

The Generalized Degrees of Freedom Region of the MIMO Z-Interference Channel with Delayed CSIT

Kaniska Mohanty and Mahesh K. Varanasi

Abstract

The generalized degrees of freedom (GDoF) region of the multiple-input multiple-output (MIMO) Gaussian Z-interference channel with an arbitrary number of antennas at each node is established under the assumption of delayed channel state information at transmitters (CSIT). The GDoF region is parameterized by α , which links the interference-to-noise ratio (INR) to the signal-to-noise ratio (SNR) via $\text{INR} = \text{SNR}^\alpha$. A new outer bound for the GDoF region is established by maximizing a bound on the weighted sum-rate of the two users, which in turn is obtained by using a combination of genie-aided side-information and an extremal inequality. The maximum weighted sum-rate in the high SNR regime is shown to occur when the transmission covariance matrix of the interfering transmitter has full rank. An achievability scheme based on block-Markov encoding and backward decoding is developed which uses interference quantization and digital multicasting to take advantage of the channel statistics of the cross-link, and the scheme is separately shown to be GDoF-optimal in both the weak ($\alpha \leq 1$) and strong ($\alpha > 1$) interference regimes. This is the first complete characterization of the GDoF region of any interference network with delayed CSIT, as well as the first such GDoF characterization of a MIMO network with delayed CSIT and arbitrary number of antennas at each node. For all antenna tuples, the GDoF region is shown to be equal to or larger than the degrees of freedom (DoF) region over the entire range of α , which leads to a V-shaped maximum sum-GDoF as a function of α , with the minimum occurring at $\alpha = 1$. The delayed CSIT GDoF region and the sum-DoF are compared with their counterparts under perfect CSIT, thereby characterizing all antenna tuples and ranges of α for which delayed CSIT is sufficient to achieve the perfect CSIT GDoF region (or sum-DoF). It is also shown that treating interference as noise is not, in general, GDoF-optimal for the MIMO Z-IC, even in the weak interference regime.

Index Terms

Channel state information, Delayed CSIT, Generalized degrees of freedom, MIMO, Z-interference channel.

I. INTRODUCTION

THE Z-interference channel (Z-IC) models two transmitter/receiver pairs communicating over a shared medium, such that only one transmitter causes interference at its unpaired receiver. It is hence the simplest model that incorporates all of the following salient features of wireless communications: broadcast, superposition, distributed transmission and links with disparate strengths. A study of the Z-IC is therefore tantamount to the study of the simplest setting in which these features are all simultaneously present in a wireless network. From a GDoF perspective, the Z-IC can be used to model a two-user interference channel (IC) in which the interference strength over one of the cross links is sufficiently weak compared to that of the other links that it can be mathematically modeled as the regime in which the high-SNR limit of the ratio of INR to SNR, each measured in the dB scale, is zero.

The delayed CSIT model was introduced for the K -user multiple-input single-output (MISO) broadcast channel (BC) in [1], where it was shown that there is a significant DoF advantage (over the no CSIT scenario) when the transmitter knows the previous channel states, even when these states are independent of the unknown current channel state. Delayed CSIT can arise in mobile communication scenarios with short coherence blocks (caused by a rapidly varying mobile environment) where CSIT through feedback links can be outdated. So far, research on delayed CSIT has mainly focused on the *DoF* region of various networks, in which all relevant communication links are implicitly assumed to be statistically comparable in strength. This includes the characterization of the DoF regions of the two-user MIMO BC, Z-IC and IC with delayed CSIT in [2], [3] and [4], respectively. Achievable DoF results (without converses) for SISO interference and X networks with delayed CSIT were obtained in [5], and bounds on the DoF for the symmetric MIMO three-user BC were obtained in [6]. Under the constraint of linear encoding strategies with delayed CSIT, the sum-DoF of the two-user MIMO X-channel was characterized in [7], as was the DoF region of the two-user symmetric MIMO X-channel.

In practical wireless networks however, signal strengths are disparate in general. In such settings, the GDoF metric is eminently more suitable. Introduced in the context of the single-input single-output (SISO) IC in [8], GDoF measures the rate of linear growth of the capacity region relative to a nominal log SNR with increasing nominal SNR, when the SNRs and INRs are assigned different *exponents* with respect to the nominal SNR. The GDoF region of the SISO IC was found in [8], a result that was later generalized in [9], to the MIMO IC with an arbitrary number of antennas at each node. Both these works assumed perfect and instantaneous CSIT.

The first step towards studying GDoF under channel uncertainty was taken in [10], wherein the GDoF region of the MIMO IC with no CSIT was characterized for the weak interference regime under certain antenna configurations. Later, [11] showed that, in the slow-fading MIMO IC, the generalized multiplexing gain region of the MIMO IC with perfect CSIT could be achieved by using *quantized* CSIT, with a sufficiently fast scaling of the number of feedback bits with the INR.

More recently, [12] studied the GDoF region of the two-user MISO BC with fixed and alternating topologies (link strengths) under alternating CSIT, where the CSIT state of each link can be either perfect and instantaneous (P), delayed (D) or not known (N). Inner and outer bounds on the GDoF under delayed CSIT were obtained therein, which are, however, not tight, so that the characterization of the GDoF region of even a two-user MISO BC with a static topology and delayed CSIT remains an open problem.

The contributions of this paper can be summarized as follows. The GDoF region of the MIMO Z-IC under the assumption of delayed CSIT is established, providing the first complete GDoF characterization of a *MIMO* network under the delayed CSIT assumption (and with an arbitrary number of antennas at each node) and also the first one obtained for an *interference* network (with distributed transmitters) and delayed CSIT. By using a genie-aided technique from [4] to provide side-information to one of the receivers and by applying the extremal inequality from [13], we bound a weighted sum-rate of the two users from above. We prove that, for both weak interference ($\alpha \leq 1$) and strong interference ($\alpha > 1$), (with $\alpha = \frac{\log \text{INR}}{\log \text{SNR}}$), the asymptotic approximation of this bound in the high SNR regime is maximized when the involved covariance matrix has full rank, thereby obtaining a new outer bound for the GDoF

region. We also design a new achievability scheme with block-Markov encoding and backward decoding that is able to achieve this outer bound in both the weak and strong interference regimes. The achievability scheme uses interference quantization in each block to compress the previous block's interference, which in turn is reconstructed using delayed CSIT. It then digitally multicasts the quantization index of the previous interference as a common message for both receivers. Decoding starts from the final block, proceeding successively backwards to the first block, using the common message decoded in block $b + 1$ as side-information while decoding block b . By specifying the GDoF carried by the common and private messages, as well as the transmit power at each transmitter, all the corner points of the GDoF outer bound region in both the strong and weak interference regimes are shown to be achievable. The GDoF region is found to be equal to or larger than the corresponding DoF region ($\alpha = 1$) for all antenna tuples in both the weak and strong interference regimes, thus demonstrating the advantage of incorporating knowledge of disparity in link strengths under delayed CSIT. The sum-GDoF, as a function of α , is shown to be V-shaped, with the minimum occurring at $\alpha = 1$. Moreover, by comparing the delayed CSIT GDoF region and sum-DoF with their perfect CSIT counterparts, we characterize all the antenna-tuples and ranges of α where delayed CSIT is sufficient to attain the perfect CSIT GDoF region and sum-GDoF. We also illustrate, with an example, that treating interference as noise is not always a GDoF-optimal strategy for the MIMO Z-IC with delayed CSIT.

It is notable that the quantization and multicasting of the interference were also used in DoF-optimal schemes for the so-called mixed CSIT model, where each transmitter has access to an imperfect estimate of the current channel in addition to accurate delayed CSIT. This commonality is not surprising however, since mixed CSIT achievability schemes, like GDoF schemes, must manage interference in the signal power level dimension. In the case of mixed CSIT, these issues arise because transmit beamforming in the null space of the (imperfect) current channel estimate results in different power levels for the interference and the signal at a receiver, a situation that also arises in the GDoF model because of the inherent disparity in the strengths of different links.

In particular, [14] presents an achievability scheme under the mixed CSIT model with features similar to ours, e.g., block-Markov encoding, introduced originally in [15] for the relay channel and later for channels involving feedback in [16] and [17], and backward decoding and interference quantization, also used in earlier works on mixed CSIT, e.g., [18], [19] and [20].

The GDoF-optimal scheme presented in this work under delayed CSIT, however, differs from the scheme in [14] in several ways. Unlike the mixed CSIT model, where the interference can only be attenuated (through beamforming at the transmitters), the GDoF model allows for interference that is stronger than the desired signal, when $\alpha > 1$. Similarly, the common message in each block, which carries the quantization index of the compressed interference from the previous block, is always received at a fixed power level in the achievable scheme from [14], but varies with α in the GDoF achievable scheme. Thus, it is clear that incorporating the channel statistics of the interfering link in a GDoF-optimal achievable scheme requires a different analysis of the achievable GDoF region compared to the mixed CSIT DoF region. Furthermore, the GDoF regions in the weak and strong interference regimes have different corner points and need to be considered separately, and require separate transmit power and common message rate allocations to achieve those corner points. Moreover, absence of any current channel estimate in the delayed CSIT model precludes any kind of transmitter beamforming, a significant feature of the mixed CSIT achievability schemes. Also, in a MIMO system, transmit beamforming with mixed CSIT can attenuate the power level of only those data streams that are in the null space of the channel estimate of a cross-link (when such a null space exists), while in the delayed CSIT GDoF model the power level of *all* data streams received over the cross-link are equally affected, and are even strengthened when $\alpha > 1$. Other related work on mixed CSIT includes the characterization of the DoF region of the MISO BC in [18], [19] and [20], and the MIMO Z-IC in [21].

The rest of the paper is organized as follows. The channel model is described in the next section. The main result of this paper is presented in Section III, followed by the proof of the outer bound in Section IV. In Section V, we describe the general achievability scheme, which is shown to be GDoF optimal for

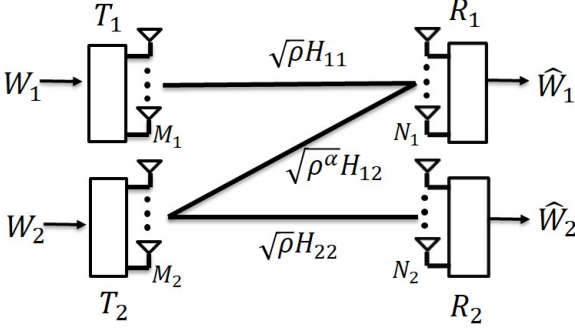


Fig. 1: The (M_1, M_2, N_1, N_2) MIMO Z-IC.

both the weak interference and strong interference regimes, in Sections VI and VII, respectively. Various aspects of the main result are discussed in Section VIII, and we conclude the paper in Section IX.

The notation used in this paper is as follows: \mathbb{R} and \mathbb{Z}^+ refer to the set of real numbers and non-negative integers, respectively. Logarithm to the base 2 is denoted by $\log(\cdot)$. The conjugate transpose of a matrix A is denoted as A^\dagger , its determinant as $|A|$ and the trace of A as $\text{tr}(A)$. \mathbf{I}_n is the identity matrix of size $n \times n$. For two matrices A and B , $A \preceq B$ means that the matrix $B - A$ is positive semi-definite. $(x)^+$ refers to the maximum of a real number x and 0. $\mathcal{CN}(0, Q)$ refers to the distribution of complex circularly symmetric Gaussian random vector with zero mean and covariance matrix Q . $\mathbb{E}[X]$ is the expectation of a random variable X . We use the standard Landau notation, where $\mathcal{O}(1)$ refers to any quantity that is bounded above by a constant. The approximation $g(\rho) \sim h(\rho)$ is a shorthand for $\lim_{\rho \rightarrow \infty} \frac{g(\rho)}{h(\rho)} = C$, where C is a constant that does not scale with ρ .

II. THE CHANNEL MODEL

The (M_1, M_2, N_1, N_2) MIMO Z-IC consists of two transmitters T_1 and T_2 with M_1 and M_2 antennas, respectively, and their paired receivers R_1 and R_2 , with N_1 and N_2 antennas, respectively. Each transmitter T_i sends a unicast message W_i to its paired receiver R_i , $i \in \{1, 2\}$. T_2 causes interference at R_1 , but T_1 does not cause any interference at R_2 (see Fig. 1). The received signals at the two receivers at time t are as follows:

$$\begin{aligned} Y_1(t) &= \sqrt{\rho}H_{11}(t)X_1(t) + \sqrt{\rho^\alpha}H_{12}(t)X_2(t) + Z_1(t), \\ Y_2(t) &= \sqrt{\rho}H_{22}(t)X_2(t) + Z_2(t), \end{aligned}$$

where $X_i(t) \in \mathbb{C}^{M_i \times 1}$ is the transmitted signal from T_i , $Y_i(t) \in \mathbb{C}^{N_i \times 1}$ is the received signal at R_i , $H_{ji}(t) \in \mathbb{C}^{N_j \times M_i}$ is the channel matrix from T_i to R_j , $Z_i(t) \sim \mathcal{CN}(0, \mathbf{I}_{N_i})$ is the additive Gaussian noise (with unit variance) at R_i , and ρ and ρ^α , where $\rho > 0$ and $\alpha \geq 0$, are the channel gains of the direct links and interfering link, respectively, for $i \in \{1, 2\}$. Transmitter T_i has an average power constraint $\text{tr}(Q_i) \leq 1$, where $Q_i \triangleq \mathbb{E}(X_i X_i^\dagger)$, $i \in \{1, 2\}$. All entries of all channel matrices are independent and identically distributed (i.i.d.), and the channel matrices and noise are assumed to be i.i.d. complex Gaussian (with unit variance) across time and independent of each other. Thus, the INR at R_1 is ρ^α while the SNR at both receivers is ρ . We define the interference to be weak when $\alpha \leq 1$, and the interference is said to be strong when $\alpha > 1$. We also define $\mathcal{H}(t) \triangleq \{H_{11}(t), H_{12}(t), H_{22}(t)\}$, and all channel matrices up to time τ are denoted by $\mathcal{H}^\tau \triangleq \{\mathcal{H}(t)\}_{t=1}^\tau$.

Both receivers have perfect knowledge of all the channel matrices. In other words, the decoding function at receiver R_i , $\forall i \in \{1, 2\}$, for a codeword spanning n channel uses is $g_i(\{Y_i(t)\}_{t=1}^n, \mathcal{H}^n) = \hat{W}_i$, where \hat{W}_i is the decoded message at R_i . The transmitters, on the other hand, learn the channel matrices only after a finite delay which, without loss of generality, we assume to be 1. Thus, at time t , each transmitter knows

all the channel matrices up to time $t - 1$. This is known as the delayed CSIT assumption. Consequently, the encoding function for T_i , $i \in \{1, 2\}$, at time t is $h_{i,t}(W_i, \mathcal{H}^{t-1})$.

The rate tuple $(\bar{R}_1(\rho, \alpha), \bar{R}_2(\rho, \alpha))$, where $\bar{R}_i = \frac{\log |\mathcal{W}_i|}{n}$ and $|\mathcal{W}_i|$ is the cardinality of the message set \mathcal{W}_i at T_i , is said to be achievable if there exists a codeword spanning n channel uses such that the probability of error goes to zero as $n \rightarrow \infty$. The capacity region $C(\rho, \alpha)$ is the region of all such achievable rate tuples, and the GDoF region is defined as the pre-log factor of the capacity region as $\rho \rightarrow \infty$, i.e.,

$$\mathbf{D} = \left\{ (d_1, d_2) \left| d_i \geq 0 \text{ and } \exists (\bar{R}_1(\rho, \alpha), \bar{R}_2(\rho, \alpha)) \in C(\rho, \alpha), \right. \right. \\ \left. \left. \text{such that } d_i = \lim_{\rho \rightarrow \infty} \frac{\bar{R}_i(\rho, \alpha)}{\log(\rho)}, i \in \{1, 2\} \right\}.$$

We also define the sum-GDoF as follows:

$$d_{\Sigma} = \sup \{d_1 + d_2 \mid (d_1, d_2) \in \mathbf{D}\}.$$

Note that in our notation R_j denotes the j^{th} receiver and \bar{R}_j denotes the j^{th} user's information rate.

III. MAIN RESULT

In the following lemma, borrowed from [9], we define a function $f()$ which provides an approximation (up to an $\mathcal{O}(1)$ term) of the sum-rate upper bound of the 2-user MIMO multiple-access channel (MAC) in the asymptotically high SNR regime. The function $f()$ will be useful not only in stating the GDoF region of the MIMO Z-IC in Theorem 1, but also in obtaining asymptotic approximations throughout the paper.

Lemma 1. *Let $H_1 \in \mathbb{C}^{u \times u_1}$ and $H_2 \in \mathbb{C}^{u \times u_2}$ be two full rank (with probability 1) channel matrices such that the matrix $H \triangleq [H_1 \ H_2]$ is also full rank (with probability 1). Then, for $\rho \rightarrow \infty$, we have*

$$\log \left| \mathbf{I}_u + \rho^{a_1} H_1 H_1^\dagger + \rho^{a_2} H_2 H_2^\dagger \right| \\ = f(u, (a_1, u_1), (a_2, u_2)) \log(\rho) + \mathcal{O}(1),$$

where for any $(u, u_1, u_2) \in \mathbb{Z}^{+3}$ and $(a_1, a_2) \in \mathbb{R}^2$, the function f is defined as

$$f(u, (a_1, u_1), (a_2, u_2)) = \min(u, u_{i_1}) a_{i_1}^+ \\ + \min((u - u_{i_1})^+, u_{i_2}) a_{i_2}^+,$$

for $i_1 \neq i_2 \in \{1, 2\}$ such that $a_{i_1} \geq a_{i_2}$.

The next theorem states the main result of this paper.

Theorem 1. *The GDoF region of the MIMO Z-IC with delayed CSIT is given by the following set of inequalities:*

$$d_1 \leq \min(M_1, N_1), \quad d_2 \leq \min(M_2, N_2), \quad (1)$$

$$\frac{d_1}{\min(M_2, N_1)} + \frac{d_2}{\min(M_2, N_1 + N_2)} \leq \frac{f(N_1, (\alpha, M_2), (1, M_1))}{\min(M_2, N_1)} + \\ + \frac{f(M_2, (\alpha, N_1), (1, N_2))}{\min(M_2, N_1 + N_2)} - \alpha. \quad (2)$$

Proof: The inequalities in (1) are the single-user outer bounds for the individual MIMO point-to-point channels. The outer bound (2) is proved in the next section.

The general achievability scheme is described in Section V, and it is shown to achieve the GDoF outer bound region in both the weak and strong interference regimes, in Sections VI and VII, respectively. ■

When $M_2 > N_1 + N_2$, it is clear that the GDoF region in Theorem 1 does not depend on M_2 . In this case, without loss of generality, we can switch off the extra transmit antennas at T_2 and prove achievability of the GDoF region using the general achievability scheme from Section V with $M_2 = N_1 + N_2$. Thus, from an achievability point of view, we can always assume $M_2 \leq N_1 + N_2$. Using similar arguments, the general achievability scheme in Section V can be restricted (without loss of generality) to the following antenna configurations:

$$\begin{aligned} M_1 &\leq N_1, & N_1 &\leq M_1 + M_2, \\ N_2 &\leq M_2, & M_2 &\leq N_1 + N_2. \end{aligned} \quad (3)$$

Define $N'_1 \triangleq \min(M_2, N_1)$. The following corollary specifies the sum-GDoF.

Corollary 1. *Under the antenna assumptions (3), the sum-GDoF of the MIMO Z-IC with delayed CSIT when $\alpha \leq 1$ is as follows:*

$$d_\Sigma = \min\left(M_1 + N_2, M_1 + N_2 - \alpha \left(M_1 - N_1 + \frac{N_2 N'_1}{M_2}\right)\right), \quad (4)$$

and the sum-GDoF when $\alpha > 1$ is as follows:

$$d_\Sigma = \min\left(M_1 + N_2, N_2 + N_1 - \frac{(N_2 + N'_1) N'_1}{M_2} + \frac{(N'_1)^2}{M_2} \alpha\right). \quad (5)$$

Proof: In Sections VI and VII, we show that (4) and (5), respectively, are direct consequences of Theorem 1 and the antenna assumptions in (3), when $\alpha \leq 1$ and $\alpha > 1$, respectively. ■

IV. PROOF OF OUTER BOUND

To prove outer bound (2), we begin by using the genie-aided technique from [4] and [22], where side-information about R_1 's message and received signal is provided to R_2 . A physically degraded channel is thus obtained, and Fano's inequality is used to bound the individual rate of each user. A weighted sum of these individual rates is, in turn, bounded using the extremal inequality from [13] for physically degraded channels, by following the steps in [14] and [18]. Subsequently, we show that the asymptotic approximation of this weighted sum-rate bound at high SNR, obtained by applying Lemma 1, is maximized, for all antenna configurations and values of α , when the transmit covariance matrix at T_2 is full rank. This maximization yields the delayed CSIT GDoF outer bound (2).

We define the following virtual received signals, obtained by subtracting the effect of X_1 from the received signal at each receiver:

$$\begin{aligned} \bar{Y}_1(t) &\triangleq \sqrt{\rho^\alpha} H_{12}(t) X_2(t) + Z_1(t), \\ \bar{Y}_2(t) &\triangleq \sqrt{\rho} H_{22}(t) X_2(t) + Z_2(t), \end{aligned}$$

and introduce the notation $\bar{Y}_i^k \triangleq \{\bar{Y}_i(t)\}_{t=1}^k$, $Y_i^k \triangleq \{Y_i(t)\}_{t=1}^k$ and $X_i^k \triangleq \{X_i(t)\}_{t=1}^k$, $\forall i \in \{1, 2\}$. Note that $Y_2(t)$ and $\bar{Y}_2(t)$ are exactly the same, but we use the above notation anyway, for convenience. Since the probability of error $P_e^{(n)}$ goes to zero as $n \rightarrow \infty$, we denote $n\epsilon_n \triangleq 1 + n\bar{R}P_e^{(n)}$ so that $\lim_{n \rightarrow \infty} \epsilon_n = 0$.

We next bound the achievable rate for the first user as follows:

$$\begin{aligned}
& n (\bar{R}_1 - \epsilon_n) \\
& \stackrel{(a)}{\leq} I(W_1; Y_1^n | \mathcal{H}^n) \\
& = I(W_1, W_2; Y_1^n | \mathcal{H}^n) - I(W_2; Y_1^n | W_1, \mathcal{H}^n) \\
& \stackrel{(b)}{=} h(Y_1^n | \mathcal{H}^n) - I(W_2; Y_1^n | W_1, \mathcal{H}^n) + n\mathcal{O}(1) \\
& \stackrel{(c)}{\leq} \sum_{t=1}^n h(Y_1(t) | \mathcal{H}(t)) - h(Y_1^n | W_1, \mathcal{H}^n) + n\mathcal{O}(1) \\
& \stackrel{(d)}{=} \sum_{t=1}^n h(Y_1(t) | \mathcal{H}(t)) - h(\bar{Y}_1^n | \mathcal{H}^n) + n\mathcal{O}(1) \\
& \stackrel{(e)}{\leq} \sum_{t=1}^n h(Y_1(t) | \mathcal{H}(t)) \\
& \quad - \sum_{t=1}^n h(\bar{Y}_1(t) | \mathcal{H}^n, \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}) + n\mathcal{O}(1) \\
& \stackrel{(f)}{=} \sum_{t=1}^n h(Y_1(t) | \mathcal{H}(t)) \\
& \quad - \sum_{t=1}^n h(\bar{Y}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) + n\mathcal{O}(1), \tag{6}
\end{aligned}$$

where (a) follows from Fano's inequality, (b) and (c) are true because Y_1^n is a deterministic function (up to $\mathcal{O}(1)$ approximation) of W_1 , W_2 and \mathcal{H}^n . We obtain (d) by removing the effect of W_1 from Y_1^n to obtain the virtual signal \bar{Y}_1^n , and (e) uses the fact that conditioning decreases entropy. Finally, we define $\mathcal{U}(t) \triangleq \{\bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^{t-1}\}$ and the equality (f) holds because, given $\mathcal{U}(t)$ and $\mathcal{H}(t)$, $\bar{Y}_1(t)$ is independent of $\{\mathcal{H}(\tau)\}_{\tau=t+1}^n$.

Next, as seen in [4] and [22], a genie provides receiver R_2 with both R_1 's message W_1 and the received

signals Y_1^n at R_1 . Now, using Fano's inequality again, the achievable rate for user 2 is bounded as follows:

$$\begin{aligned}
& n (\bar{R}_2 - \epsilon_n) \\
& \leq I(W_2; Y_1^n, Y_2^n, W_1 | \mathcal{H}^n) \\
& \stackrel{(a)}{=} I(W_2; Y_1^n, Y_2^n | W_1, \mathcal{H}^n) \\
& \stackrel{(b)}{=} I(W_2; \bar{Y}_1^n, \bar{Y}_2^n | \mathcal{H}^n) \\
& = \sum_{t=1}^n I(W_2; \bar{Y}_1(t), \bar{Y}_2(t) | \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^n) \\
& \stackrel{(c)}{\leq} \sum_{t=1}^n I(X_2(t); \bar{Y}_1(t), \bar{Y}_2(t) | \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^n) \\
& = \sum_{t=1}^n (h(\bar{Y}_1(t), \bar{Y}_2(t) | \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^n) \\
& \quad - h(\bar{Y}_1(t), \bar{Y}_2(t) | X_2(t), \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^n)) \\
& \leq \sum_{t=1}^n h(\bar{Y}_1(t), \bar{Y}_2(t) | \bar{Y}_1^{t-1}, \bar{Y}_2^{t-1}, \mathcal{H}^n) \\
& \stackrel{(d)}{=} \sum_{t=1}^n h(\bar{Y}_1(t), \bar{Y}_2(t) | \mathcal{U}(t), \mathcal{H}(t)), \tag{7}
\end{aligned}$$

where (a) is true because W_1 and W_2 are independent of each other, (b) is obtained by removing the effect of W_1 from the received signals Y_1^n and Y_2^n , to obtain the virtual received signals \bar{Y}_1^n and \bar{Y}_2^n , respectively, and the inequality in (c) is an application of the data processing inequality. Finally, (d) is true because, given $\mathcal{U}(t)$ and $\mathcal{H}(t)$, both $\bar{Y}_1(t)$ and $\bar{Y}_2(t)$ are independent of $\{\mathcal{H}(\tau)\}_{\tau=t+1}^n$.

We now define the following:

$$\begin{aligned}
S(t) & \triangleq \begin{bmatrix} \sqrt{\rho^\alpha} H_{12}(t) \\ \sqrt{\rho} H_{22}(t) \end{bmatrix}, \\
K(t) & \triangleq \mathbb{E} \left(x_2(t) x_2^\dagger(t) | \mathcal{U}(t) \right), \\
L(t) & \triangleq \mathbb{E} \left(x_1(t) x_1^\dagger(t) | \mathcal{H}^{t-1} \right), \\
\mathcal{V}(t) & \triangleq \{ \mathcal{U}(t), \mathcal{H}(t) \}, \\
p & \triangleq \min(M_2, N_1 + N_2), \\
q & \triangleq \min(M_2, N_1) \triangleq N'_1. \tag{8}
\end{aligned}$$

As shown in [14] and [18], by applying the extremal inequality for degraded outputs (see [13], [23]) to the physically degraded channel $X_2^n \rightarrow (\bar{Y}_1^n, \bar{Y}_2^n) \rightarrow \bar{Y}_1^n$, we obtain the following inequality:

$$\begin{aligned}
& \frac{1}{p} h(\bar{Y}_1(t), \bar{Y}_2(t) | \mathcal{V}(t)) - \frac{1}{q} h(\bar{Y}_1(t) | \mathcal{V}(t)) \\
& \leq \max_{K \succeq 0} \mathbb{E}_S \left(\frac{1}{p} \log |\mathbf{I}_{N_1+N_2} + S(t) K(t) S^\dagger(t)| \right. \\
& \quad \left. \text{tr}(K) \leq 1 \right. \\
& \quad \left. - \frac{1}{q} \log |\mathbf{I}_{N_1} + \rho^\alpha H_{12}(t) K(t) H_{12}^\dagger(t)| \right), \tag{9}
\end{aligned}$$

which allows us to outer bound the weighted sum of the achievable rate of the two users from (6) and (7). But first, we use Lemma 1 to obtain the asymptotic approximation in the high SNR regime of each of the terms in (9).

Note that, because of delayed CSIT, the covariance matrix $K(t)$ is independent of both $H_{12}(t)$ and $H_{22}(t)$, and is thus also independent of $S(t)$. Now, let the matrix $K(t)$ have rank $M_2 - r$, such that $0 \leq r \leq M_2$ with $r = 0$ corresponding to a full rank $K(t)$. The singular value decomposition (SVD) of the transmit covariance matrix $K(t)$ can be written as $K(t) = U(t) \Lambda(t) U^\dagger(t)$, where $U(t) \in \mathbb{C}^{M_2 \times (M_2 - r)}$ is such that $U^\dagger U = \mathbf{I}$ and $\Lambda(t)$ is a $(M_2 - r) \times (M_2 - r)$ diagonal matrix containing the non-zero singular values of $K(t)$ in descending order. Using the SVD of $K(t)$, we obtain the following asymptotic approximation of the first term in (9) (we suppress the index t henceforth):

$$\begin{aligned} & \log |\mathbf{I}_{N_1+N_2} + SKS^\dagger| \\ & \stackrel{(a)}{=} \log \left| \mathbf{I}_{N_1+N_2} + \begin{bmatrix} \sqrt{\rho^\alpha} \tilde{H}_{12} \\ \sqrt{\rho} \tilde{H}_{22} \end{bmatrix} \begin{bmatrix} \sqrt{\rho^\alpha} \tilde{H}_{12}^\dagger \\ \sqrt{\rho} \tilde{H}_{22}^\dagger \end{bmatrix}^\dagger \right| \\ & \stackrel{(b)}{=} \log |\mathbf{I}_{M_2-r} + \rho^\alpha \tilde{H}_{12}^\dagger \tilde{H}_{12} + \rho \tilde{H}_{22}^\dagger \tilde{H}_{22}| \\ & = f(M_2 - r, (\alpha, N_1), (1, N_2)) \log \rho + n\mathcal{O}(1), \end{aligned} \quad (10)$$

where, in (a), we have defined $\tilde{H}_{i2} \triangleq H_{i2} U \Lambda^{\frac{1}{2}}$, for $i \in \{1, 2\}$, and in (b), we use the identity $|\mathbf{I} + AB| = |\mathbf{I} + BA|$ and finally, (10) follows from Lemma 1.

Similarly, using the SVD of $K(t)$ again, we obtain the following asymptotic approximation of the second term in (9):

$$\begin{aligned} & \log |\mathbf{I}_{N_1} + \rho^\alpha H_{12} K H_{12}^\dagger| \\ & = \alpha \min(M_2 - r, N_1) \log \rho + n\mathcal{O}(1). \end{aligned} \quad (11)$$

Next, we approximate the first term in (6) as follows:

$$\begin{aligned} & \sum_{t=1}^n h(Y_1(t) | \mathcal{H}(t)) \\ & \stackrel{(a)}{\leq} n \log |\mathbf{I}_{N_1} + \rho H_{11} L H_{11}^\dagger + \rho^\alpha H_{12} K H_{12}^\dagger| \\ & \stackrel{(b)}{\leq} n \log |\mathbf{I}_{N_1} + \rho H_{11} H_{11}^\dagger + \rho^\alpha \tilde{H}_{12} \tilde{H}_{12}^\dagger| \\ & \stackrel{(c)}{=} n f(N_1, (1, M_1), (\alpha, M_2 - r)) \log \rho + n\mathcal{O}(1), \end{aligned} \quad (12)$$

where (a) uses the fact that Gaussian inputs maximize the entropy. Since $\text{tr}(L) \leq 1$, we have $L \preceq \mathbf{I}_{M_1}$ and (b) follows by using the SVD of K and substituting L with \mathbf{I}_{M_1} , since $\log \det$ is monotonically increasing on the cone of positive definite matrices. The equality in (c) is a direct consequence of Lemma 1.

Finally, we outer bound the weighted sum of the achievable rates from (6) and (7) as follows:

$$\begin{aligned}
& n \left(\frac{\bar{R}_1}{q} + \frac{\bar{R}_2}{p} - \epsilon_n \right) \\
& \stackrel{(a)}{\leq} n \cdot \frac{1}{q} f(N_1, (1, M_1), (\alpha, M_2 - r)) \log \rho + n \mathcal{O}(1) \\
& \quad + \frac{1}{p} \sum_{t=1}^n h(\bar{Y}_1(t), \bar{Y}_2(t) | \mathcal{V}(t)) - \frac{1}{q} \sum_{t=1}^n h(\bar{Y}_1(t) | \mathcal{V}(t)) \\
& \stackrel{(b)}{\leq} \frac{n}{q} f(N_1, (1, M_1), (\alpha, M_2 - r)) \log \rho \\
& \quad + n \frac{f(M_2 - r, (\alpha, N_1), (1, N_2))}{p} \log \rho \\
& \quad - n \frac{\alpha \min(M_2 - r, N_1)}{q} \log \rho + n \mathcal{O}(1) \tag{13} \\
& \stackrel{(c)}{\leq} \frac{n}{q} f(N_1, (1, M_1), (\alpha, M_2)) \log \rho \\
& \quad + n \left(\frac{f(M_2, (\alpha, N_1), (1, N_2))}{p} - \frac{\alpha N'_1}{q} \right) \log \rho + n \mathcal{O}(1), \tag{14}
\end{aligned}$$

where inequality (a) is obtained by using (7) and substituting (12) in (6), and the second inequality (b) is obtained by using (9), after substituting the asymptotic approximations from (10) and (11). As shown below in Lemma 2, the expression in (13) is maximized for both weak and strong interference when $r = 0$, and thus, we substitute $r = 0$ to obtain the inequality in (c). Finally, substituting the values of p and q from (8) in (14) and dividing both sides of the inequality by $n \cdot \log \rho$ as $\rho \rightarrow \infty$ (for which $\epsilon_n \rightarrow 0$), we obtain the outer bound (2).

Lemma 2. *The function $g(r)$, defined as follows,*

$$\begin{aligned}
g(r) \triangleq & \frac{f(N_1, (1, M_1), (\alpha, M_2 - r))}{\min(M_2, N_1)} \\
& + \frac{f(M_2 - r, (\alpha, N_1), (1, N_2))}{\min(M_2, N_1 + N_2)} \\
& - \alpha \frac{\min(M_2 - r, N_1)}{\min(M_2, N_1)}, \tag{15}
\end{aligned}$$

for $0 \leq r \leq M_2$ and $r \in \mathbb{Z}$, is maximized at $r = 0$, for $\alpha \in [0, \infty)$.

Proof: See Appendix A. ■

V. GENERAL ACHIEVABILITY SCHEME

As explained in Section III, we can restrict the achievability scheme, without any loss of generality, to the antenna configurations shown in (3). Henceforth, we operate under the antenna assumptions in (3) for the rest of the paper.

Summary of achievable scheme:

The general achievability scheme has a block-Markov structure, which consists of B blocks, each consisting of s time slots. In this paper, without loss of generality, we take $s = 1$, i.e., each block consists of a single time slot. In block b , each transmitter T_i , $i \in \{1, 2\}$ transmits a message $w_{i,b}$ intended for R_i using M_i data streams. While T_1 transmits its message at full power, T_2 modulates its transmit power

level, parameterized by A_2 , based on the GDoF tuple to be achieved. The interference seen at R_1 in the previous block $b - 1$, which has a power level $\rho^{(\alpha - A_2)}$, can be reconstructed at T_2 in this block, using delayed CSIT of the cross-link H_{12} . T_2 uses digital quantization to compress this interference, such that the average distortion does not exceed the noise level (which can then be ignored from a GDoF perspective) and digitally multicasts the quantization index l_{b-1} with full power, after encoding it as a common message $x_{2c}(l_{b-1})$. This common message is useful at both the receivers while decoding block $b - 1$, since it provides R_1 with enough information to subtract an estimate of the interference seen in block $b - 1$, and provides R_2 useful side-information about its own message $w_{2,b-1}$. Decoding starts from the final block, proceeding successively backwards to the first block, using the common message decoded in block b as side-information while decoding block $b - 1$. We thus obtain a general achievability region, parameterized by the power level A_2 . This allows us to specify the transmit power level A_2 required to achieve each non-trivial corner point of the GDoF region from Theorem 1, separately for weak and strong interference in the next two sections, respectively.

Because of the similar structure of our achievability scheme with the DoF achievability scheme for the IC with mixed CSIT in [14], we adopt the notation from [14] for presentation purposes. But, as explained earlier in the introduction, our subsequent analysis differs considerably from the DoF analysis in [14].

Encoding and transmission strategy:

In each block b , transmitter T_i , $i \in \{1, 2\}$, encodes its private message $w_{i,b}$ as the vector $u_i(w_{i,b}) \in \mathbb{C}^{M_i \times 1}$, such that $u_i(w_{i,b}) \sim \mathcal{CN}(0, Q_i)$, where $Q_1 \triangleq \mathbf{I}_{M_1}$ and $Q_2 \triangleq \rho^{-A_2} \mathbf{I}_{M_2}$, with $A_2 \geq 0$. We note that, by using the \sim notation, we can omit a constant multiplicative factor for the covariance matrices which does not affect the GDoF analysis. In other words, while T_1 transmits its message at full power $P_{T_1} \sim \rho^0$, T_2 transmits its private message at less than full power $P_{T_2} \sim \rho^{-A_2}$. In the next two sections, the power level A_2 will be specified separately for each GDoF corner point. T_2 also encodes the common message l_{b-1} , to be defined later, using the vector $x_{2c}(l_{b-1}) \in \mathbb{C}^{M_2 \times 1}$ and transmits it with maximum power $P_c \sim \rho^0$. Thus, the transmitted signal at transmitters T_1 and T_2 in block b are, respectively as follows:

$$\begin{aligned} x_1[b] &= u_1(w_{1,b}), \\ x_2[b] &= u_2(w_{2,b}) + x_{2c}(l_{b-1}). \end{aligned}$$

Since there is no common message in the first block $b = 1$, we set $l_0 = 0$. In the final block $b = B$, only the common message is transmitted, and so we set $w_{1,B} = w_{2,B} = 0$, to end the transmission.

In block b , the received signal at each receiver, and the power level of each constituent of the received signal (indicated below it), is as follows:

$$\begin{aligned} Y_1[b] &= \underbrace{\sqrt{\rho^\alpha} H_{12}[b] x_{2c}(l_{b-1})}_{\rho^\alpha} + \underbrace{\sqrt{\rho} H_{11}[b] u_1(w_{1,b})}_{\rho^1} \\ &\quad + \underbrace{\sqrt{\rho^\alpha} H_{12}[b] u_2(w_{2,b})}_{\eta_b \sim \rho^{(\alpha - A_2)}}, \end{aligned}$$

and

$$Y_2[b] = \underbrace{\sqrt{\rho} H_{22}[b] x_{2c}(l_{b-1})}_{\rho^1} + \underbrace{\sqrt{\rho} H_{22}[b] u_2(w_{2,b})}_{\rho^{(1 - A_2)}},$$

where η_b is the interference caused at R_1 by u_2 . This interference η_b is reconstructed at T_2 in the next block $b + 1$ through delayed CSIT of the channel $H_{12}[b]$. From the rate distortion theorem [24], we know that this interference, which is at power level $\rho^{(\alpha - A_2)}$ and carries a GDoF of no more than that of u_2 , i.e., N_2 , can be quantized at T_2 using a source codebook of size $\rho^{\min(N_2, (\alpha - A_2)N'_1)}$, such that the mean square distortion does not exceed the noise level and can thus be ignored from a GDoF perspective. The

quantization index is denoted as l_b , and we also define d_η as the GDoF carried by the common message. The GDoF carried by the private message $w_{i,b}$ is denoted as d_{ib} , $\forall i \in \{1, 2\}$.

Decoding:

The decoding starts from the last block B , in which both receivers decode only the common message l_{B-1} , since $w_{1,B} = w_{2,B} = 0$. This is possible only when

$$d_\eta \leq \alpha \min(M_2, N_1) = \alpha N'_1, \quad (16)$$

$$d_\eta \leq \min(M_2, N_2), \quad (17)$$

where the two conditions are the decoding conditions at R_1 and R_2 , respectively. The first condition comes from the MIMO point-to-point channel from T_2 to R_1 , with the received power of the common message l_{B-1} being ρ^α , and similarly, the second condition comes from the MIMO point-to-point channel from T_2 to R_2 , where the message l_{B-1} is received at power ρ .

The decoding process now moves backwards to the previous block. Thus, while decoding block b , the common message l_b is known at each receiver from the decoding of block $b+1$. Receiver R_1 can thus reconstruct the interference η_b (with a distortion that does not exceed the noise level which can thus be neglected) and subtract it from its received signal, while receiver R_2 reconstructs η_b to use it as side-information. Now, the signal at each receiver is that of a 2-user MAC, as shown below (for the sake of brevity, we omit the block indices b for the channel matrices):

$$Y_1[b] - \eta_b = \sqrt{\rho^\alpha} H_{12} x_{2c}(l_{b-1}) + \sqrt{\rho} H_{11} u_1(w_{1,b}), \quad (18)$$

$$\begin{bmatrix} Y_2[b] \\ \eta_b \end{bmatrix} = \begin{bmatrix} \sqrt{\rho} H_{22} \\ 0 \end{bmatrix} x_{2c}(l_{b-1}) + \begin{bmatrix} \sqrt{\rho} H_{22} \\ \sqrt{\rho^\alpha} H_{12} \end{bmatrix} u_2(w_{2,b}). \quad (19)$$

In Appendix B, we prove that each receiver R_i , $i \in \{1, 2\}$ can decode its own message $w_{i,b}$, carrying d_{ib} GDoF, as well as the common message l_{b-1} , which carries d_η GDoF, if the (d_{1b}, d_{2b}, d_η) tuple lies within the following achievable region:

$$d_\eta \leq \min(\alpha N'_1, N_2), \quad (20)$$

$$d_{1b} \leq M_1, \quad (21)$$

$$d_\eta + d_{1b} \leq f(N_1, (\alpha, M_2), (1, M_1)), \quad (22)$$

$$d_{2b} \leq f(M_2, (1 - A_2, N_2), (\alpha - A_2, N_1)), \quad (23)$$

$$d_\eta + d_{2b} \leq (\alpha - A_2) N'_1 + N_2. \quad (24)$$

Inequalities (20)-(24) will be referred to as the general achievability conditions. Since no private messages are transmitted in the final block, the final achievable GDoF using this achievable scheme is $d_i \triangleq \frac{1}{B} \sum_{b=1}^{B-1} d_{ib} = \frac{B-1}{B} d_{ib}$, $i \in \{1, 2\}$, where we have allocated the same d_{ib} in each block b , and by taking $B \rightarrow \infty$, we get $d_i \rightarrow d_{ib}$. Henceforth, we replace d_{1b} and d_{2b} with d_1 and d_2 , respectively. In the following two sections, we show how this general achievable scheme can be used to achieve the GDoF outer bound region in both the weak and strong interference regimes.

VI. WEAK INTERFERENCE - ACHIEVABILITY

In the weak interference regime ($\alpha \leq 1$), the general achievability conditions from the previous section can be simplified using the following lemma.

Lemma 3. *When $\alpha \leq 1$, the following (d_1, d_2) GDoF tuple can be achieved:*

$$d_1 \triangleq \min(M_1, M_1 - \alpha(M_1 + N'_1 - N_1) + N'_1 A_2), \quad (25)$$

$$d_2 \triangleq \min(N_2, N_2 + \alpha(M_2 - N_2) - M_2 A_2), \quad (26)$$

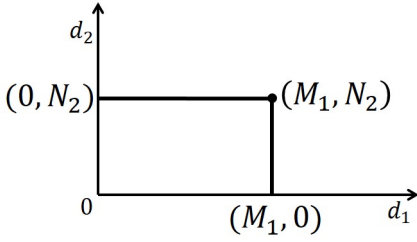


Fig. 2: GDoF region of the MIMO Z-IC with delayed CSIT, when bound (2) is inactive, as seen in Case I of Sections VI and VII.

whenever A_2 , defined in Section V, lies in the range

$$\alpha \geq A_2 \geq \left(\alpha - \frac{N_2}{N'_1} \right)^+. \quad (27)$$

Proof: To prove the lemma, we need to show that the (d_1, d_2) tuple defined above satisfies the general achievability conditions (20)-(24). We allocate

$$d_\eta \triangleq (\alpha - A_2) N'_1, \quad (28)$$

which makes the first general achievability condition (20) redundant, as long as A_2 lies in the range shown in (27). After substituting (28) in the remaining achievability conditions (21)-(24), we simplify the terms involving $f()$, by using $\alpha \leq 1$. The general achievability conditions (21) and (22) are thus respectively simplified as follows:

$$d_1 \leq M_1, \quad (29)$$

$$d_1 \leq M_1 + \alpha(N_1 - M_1) - (\alpha - A_2) N'_1. \quad (30)$$

Similarly, the simplified general achievability conditions (23) and (24) are, respectively,

$$d_2 \leq (1 - A_2) N_2 + (\alpha - A_2) (M_2 - N_2), \quad (31)$$

$$d_2 \leq N_2. \quad (32)$$

Now, by combining (29) and (30), it is clear that d_1 defined in (25) satisfies the general achievability conditions, while achievability of d_2 defined in (26) similarly follows from (31) and (32), and the lemma is thus proved. ■

Using the above lemma, we show that the outer bound region from Theorem 1 is achievable when $\alpha < 1$. In the weak interference regime, under the antenna assumptions (3), the outer bound region from Theorem 1 is given by the following set of inequalities:

$$d_1 \leq M_1, \quad (33)$$

$$d_2 \leq N_2, \quad (34)$$

$$\frac{d_1}{N'_1} + \frac{d_2}{M_2} \leq \frac{M_1 + (N_1 - M_1)\alpha}{N'_1} + \frac{N_2 + (M_2 - N_2)\alpha}{M_2} - \alpha. \quad (35)$$

This leads to two possible shapes of the outer bound region, depending on whether the delayed CSIT bound (35) is active or not. We analyze the two cases below, and show that the GDoF region is achievable in both cases.

Case I) when $\frac{N_1 - M_1}{N'_1} \geq \frac{N_2}{M_2}$:

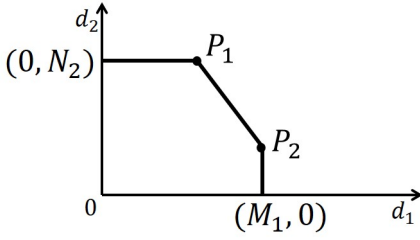


Fig. 3: GDoF region of the MIMO Z-IC with delayed CSIT, when bound (2) is active. The corner points P_1 and P_2 are defined separately for weak and strong interference in Case II of Sections VI and VII, respectively.

In this case, the delayed CSIT bound (35) is inactive, and the corresponding GDoF outer bound region is shown in Fig. 2. The only non-trivial GDoF corner point is (M_1, N_2) , which can be achieved by setting the transmission power level at T_2 to

$$A_2 = \left(1 - \frac{N_2}{M_2}\right) \alpha, \quad (36)$$

which satisfies condition (27) in Lemma 3. Substituting (36) in (25) from Lemma 3, we see that the GDoF achieved by the first user is

$$d_1 = \min \left(M_1, M_1 + \alpha N'_1 \left(\frac{N_1 - M_1}{N'_1} - \frac{N_2}{M_2} \right) \right) = M_1.$$

Similarly, substituting (36) in (26), we see that $d_2 = N_2$ is achievable, thus proving that the GDoF tuple (M_1, N_2) is achievable.

Case II) when $\frac{N_1 - M_1}{N'_1} < \frac{N_2}{M_2}$:

In this case, the delayed CSIT bound (35) is active, and the GDoF outer bound region is shown in Fig. 3, which contains two non-trivial corner points, which are listed below:

$$\begin{aligned} P_1 &\triangleq \left(M_1 - \alpha (M_1 + N'_1 - N_1) + \frac{N'_1}{M_2} (M_2 - N_2) \alpha, N_2 \right), \\ P_2 &\triangleq \left(M_1, N_2 + \alpha (M_2 - N_2) - \frac{M_2}{N'_1} (N'_1 - N_1 + M_1) \alpha \right). \end{aligned} \quad (37)$$

The first corner point P_1 is achieved by setting the transmission power level as follows:

$$A_2 = \left(1 - \frac{N_2}{M_2}\right) \alpha, \quad (38)$$

and the second corner point P_2 is achieved with the transmission power level

$$A_2 = \left(1 - \frac{N_1 - M_1}{N'_1}\right) \alpha. \quad (39)$$

For both the points, the assigned value of A_2 satisfies the condition (27) in Lemma 3, and it is straightforward to check that by substituting (38) and (39) in Lemma 3, we achieve the points P_1 and P_2 , respectively.

A. Comparison with DoF region

The DoF region of the MIMO Z-IC with delayed CSIT (with the antenna conditions (3)) was obtained in [3] and is given below:

$$\begin{aligned} d_1 &\leq M_1, \\ d_2 &\leq N_2, \\ d_1 + d_2 &\leq \max(M_2, N_1), \\ \frac{d_1}{N'_1} + \frac{d_2}{M_2} &\leq \frac{N_1}{N'_1}. \end{aligned} \quad (40)$$

Comparing it with the GDoF region with weak interference shown in (33)-(35), it is straightforward to check that the DoF delayed CSIT region (40) is always smaller than or identical to the corresponding GDoF delayed CSIT region for all values of $\alpha \leq 1$.

B. Comparison with perfect CSIT

The GDoF region of the MIMO IC with perfect CSIT was obtained in [11], from which, by setting $\alpha_{11} = \alpha_{22} = 1$, $\alpha_{12} = 0$ and $\alpha_{21} = \alpha$ in Theorem 1 therein, we obtain the GDoF region of the MIMO Z-IC with perfect CSIT, given below (under the antenna conditions (3)):

$$\begin{aligned} d_1 &\leq M_1, \quad d_2 \leq N_2, \\ d_1 + d_2 &\leq f(N_1, (\alpha, M_2), (1, M_1)) \\ &\quad + f(N_2, (1 - \alpha, N'_1), (1, M_2 - N'_1)). \end{aligned} \quad (41)$$

When $\alpha \leq 1$ and $M_2 \leq N_1$, i.e., $N'_1 = M_2$, it is straightforward to show that the GDoF region with only delayed CSIT, shown in (33)-(35), coincides with the perfect CSIT GDoF region in (41).

When $M_2 > N_1$ and $\alpha \leq 1$, the non-trivial perfect CSIT bound in (41) is always loose compared to the corresponding delayed CSIT bound (35). Thus, when $M_2 > N_1$ and the delayed CSIT bound (35) is active, i.e., when $\frac{N_1 - M_1}{N_1} < \frac{N_2}{M_2}$ (Case II above), the delayed CSIT GDoF region with weak interference is strictly smaller than the corresponding perfect CSIT GDoF region. Otherwise, for antenna configurations from Case I with $M_2 > N_1$, the delayed CSIT and perfect CSIT GDoF regions coincide for $\alpha \leq 1$.

C. Sum-GDoF

For Case I above, we see from the shape of the GDoF region in Fig. 2 that the sum-GDoF is

$$d_\Sigma = M_1 + N_2.$$

For Case II, the straight line in the (d_1, d_2) plane that defines the outer bound (2) has slope $-\frac{M_2}{N'_1} < -1$, and thus, the maximum sum-GDoF is achieved at the corner point P_1 in Fig. 3, and is equal to

$$d_\Sigma = M_1 + N_2 - \alpha \left(M_1 - N_1 + \frac{N_2 N'_1}{M_2} \right),$$

which decreases linearly with α in the weak interference regime. This proves the first part (4) of Corollary 1.

For comparison, the sum-GDoF with perfect CSIT can be obtained from (41). When $\alpha \leq 1$ and $M_2 > N_1$ (since the perfect CSIT and delayed CSIT GDoF regions are the same when $M_2 \leq N_1$), the perfect CSIT sum-GDoF is as follows:

$$d_\Sigma^{\text{p}} = \begin{cases} M_1 + N_2, & M_1 + N_2 \leq M_2 \\ M_1(1 - \alpha) + M_2\alpha + N_2(1 - \alpha), & M_1 + N_2 > M_2. \end{cases}$$

VII. STRONG INTERFERENCE - ACHIEVABILITY

When $\alpha > 1$, the general achievability conditions (20)-(24) can be simplified using the following lemma.

Lemma 4. *When $\alpha > 1$, the general achievability scheme can achieve the GDoF tuple (d_1, d_2) shown below:*

$$d_1 \triangleq \begin{cases} \min(M_1, (\alpha - 1)N'_1 + (N_1 - N_2)), & A_2 < \alpha - \frac{N_2}{N'_1} \\ \min(M_1, N_1 - N'_1 + N'_1 A_2), & A_2 \geq \alpha - \frac{N_2}{N'_1} \end{cases} \quad (42)$$

and

$$d_2 \triangleq \begin{cases} N_2, & A_2 < \alpha - \frac{N_2}{N'_1} \\ \min(N_2, (\alpha - A_2)N'_1 + (1 - A_2)^+(M_2 - N'_1)), & A_2 \geq \alpha - \frac{N_2}{N'_1} \end{cases} \quad (43)$$

where A_2 is defined in Section V.

Proof: In the strong interference regime, we assign the value of d_η as follows:

$$d_\eta \triangleq \begin{cases} N_2, & A_2 < \alpha - \frac{N_2}{N'_1} \\ (\alpha - A_2)N'_1, & A_2 \geq \alpha - \frac{N_2}{N'_1} \end{cases} \quad (44)$$

which trivially satisfies the first general achievability condition (20). Moreover, since $\alpha > 1$, the general achievability condition (22) simplifies to

$$d_\eta + d_1 \leq \alpha N'_1 + (N_1 - N'_1).$$

Substituting the value of d_η from (44) in the above condition, and combining it with the achievability condition (21), i.e., $d_1 \leq M_1$, we see that d_1 shown in (42) is always achievable.

Since $\alpha > 1$, the general achievability condition (23) simplifies to

$$d_2 \leq (\alpha - A_2)N'_1 + (1 - A_2)^+(M_2 - N'_1). \quad (45)$$

Moreover, when $A_2 \geq \alpha - \frac{N_2}{N'_1}$, we also have the condition

$$d_2 \leq N_2, \quad (46)$$

obtained by substituting $d_\eta = (\alpha - A_2)N'_1$ in the remaining general achievability condition (24). Combining (45) and (46), it is clear that when $A_2 \geq \alpha - \frac{N_2}{N'_1}$, the value of d_2 shown in (43) is achievable.

When $A_2 < \alpha - \frac{N_2}{N'_1}$, i.e., when $d_\eta = N_2$, the general achievability conditions (23) and (24) can be reframed, respectively, as (45) and

$$d_2 \leq (\alpha - A_2)N'_1,$$

both of which are satisfied by assigning $d_2 = N_2$, since we have $N_2 \leq (\alpha - A_2)N'_1$ in this case. Thus, the value of d_2 specified in the lemma is always achievable, proving the lemma. ■

Remark 1. The achievable d_2 in (43) from Lemma 4 is clearly a monotonically decreasing function of A_2 , and with some straightforward algebraic manipulation, it is easily shown that the value of d_2 in (43) remains constant at N_2 when $A_2 \leq A_2^{d2}$, and strictly monotonically decreases with A_2 when $A_2 > A_2^{d2}$, where A_2^{d2} is given below:

$$A_2^{d2} \triangleq \begin{cases} 1 + \frac{(\alpha-1)N'_1 - N_2}{M_2}, & \alpha < 1 + \frac{N_2}{N'_1} \\ \alpha - \frac{N_2}{N'_1}, & \alpha \geq 1 + \frac{N_2}{N'_1}. \end{cases} \quad (47)$$

In the strong interference regime, and using the antennas assumptions in (3), the outer bound region from Theorem 1 is as follows:

$$\begin{aligned} d_1 &\leq M_1, \\ d_2 &\leq N_2, \\ \frac{d_1}{N'_1} + \frac{d_2}{M_2} &\leq \frac{N_1}{N'_1} + \frac{(\alpha - 1) N'_1}{M_2}, \end{aligned} \quad (48)$$

where we have used the following simplifications for strong interference,

$$\begin{aligned} f(N_1, (\alpha, M_2), (1, M_1)) &= \alpha N'_1 + (N_1 - N'_1), \\ f(M_2, (\alpha, N_1), (1, N_2)) &= \alpha N'_1 + M_2 - N'_1. \end{aligned}$$

Similar to the weak interference regime, the above GDoF outer bound region can have two different shapes, as shown below, depending on whether the delayed CSIT bound (48) is active or not. We now show that the GDoF outer bound region with strong interference ($\alpha > 1$) is achievable in both cases, shown below:

Case I) when $\alpha \geq 1 + \frac{N_2}{N'_1} - \frac{M_2(N_1 - M_1)}{(N'_1)^2}$.

In this case, the delayed CSIT bound (48) is inactive, and the GDoF outer bound region is shown in Fig. 2. The only non-trivial GDoF corner point is (M_1, N_2) , which can be achieved by setting the transmission power level at T_2 as follows:

$$A_2 = 1 - \frac{N_1 - M_1}{N'_1}. \quad (49)$$

To see that the assigned value of A_2 achieves $d_2 = N_2$, it suffices to show that $A_2 \leq A_2^{d2}$, as per Remark 1. To this end, we consider both sub-cases of (47). When $\alpha < 1 + \frac{N_2}{N'_1}$, we see that $A_2 \leq A_2^{d2}$, as shown below:

$$\begin{aligned} \alpha &\geq 1 + \frac{N_2}{N'_1} - \frac{M_2(N_1 - M_1)}{(N'_1)^2} \\ \Rightarrow 1 - \frac{N_1 - M_1}{N'_1} &\leq 1 + \frac{(\alpha - 1) N'_1 - N_2}{M_2}. \end{aligned}$$

When $\alpha \geq 1 + \frac{N_2}{N'_1}$, we have

$$\begin{aligned} \alpha &\geq 1 + \frac{N_2}{N'_1} - \frac{N_1 - M_1}{N'_1} \\ \Rightarrow 1 - \frac{N_1 - M_1}{N'_1} &\leq \alpha - \frac{N_2}{N'_1}, \end{aligned}$$

i.e., $A_2 \leq A_2^{d2}$.

Substituting the value of A_2 in (42) achieves $d_1 = M_1$, as we show for each of the two sub-cases in (42). For the first sub-case, by substituting (49) in the condition $A_2 < \alpha - \frac{N_2}{N'_1}$, we obtain

$$M_1 < (\alpha - 1) N'_1 + N_1 - N_2,$$

and thus, using the above inequality in (42), it is clear that $d_1 = M_1$ is achievable. For the second sub-case, i.e., when $A_2 \geq \alpha - \frac{N_2}{N'_1}$, we see that

$$N_1 - N'_1 + N'_1 A_2 = N_1 - N'_1 + N'_1 \left(1 - \frac{N_1 - M_1}{N'_1} \right) = M_1$$

and thus $d_1 = M_1$ is achievable for this sub-case too.

Case II: when $\alpha < 1 + \frac{N_2}{N'_1} - \frac{M_2(N_1 - M_1)}{(N'_1)^2}$.

In this case, the delayed CSIT bound (48) is active, and the GDoF region is shown in Fig. 3, where the two non-trivial corner points are as follows:

$$\begin{aligned} P_1 &\triangleq \left(N_1 + \frac{N'_1}{M_2} ((\alpha - 1) N'_1 - N_2), N_2 \right), \\ P_2 &\triangleq \left(M_1, (\alpha - 1) N'_1 + \frac{M_2}{N'_1} (N_1 - M_1) \right). \end{aligned}$$

Point P_1

The corner point P_1 is achieved by setting the transmission power level as

$$A_2 = 1 - \frac{N_2 - (\alpha - 1) N'_1}{M_2}. \quad (50)$$

Since, $\alpha < 1 + \frac{N_2}{N'_1}$ (from the defining condition of Case II), we see that $A_2 = A_2^{d2}$ (from (47) and (50)) and thus $d_2 = N_2$ is achievable, as per Remark 1. As for d_1 , we see that only the second sub-case of (42) in Lemma 4 is involved, since $A_2 \leq \alpha - \frac{N_2}{N'_1}$, as seen below:

$$\begin{aligned} \alpha &< 1 + \frac{N_2}{N'_1} \\ \Rightarrow \alpha \left(1 - \frac{N'_1}{M_2} \right) &\leq \left(1 + \frac{N_2}{N'_1} \right) \left(1 - \frac{N'_1}{M_2} \right) \\ \Rightarrow 1 - \frac{N_2 - (\alpha - 1) N'_1}{M_2} &\leq \alpha - \frac{N_2}{N'_1}. \end{aligned}$$

Now, substituting the value of A_2 in the second sub-case of (42), we see that P_1 is achievable.

Point P_2

To achieve the second corner point P_2 , we set the transmission power level as

$$A_2 = 1 - \frac{N_1 - M_1}{N'_1}. \quad (51)$$

It is easy to show that $A_2 > \alpha - \frac{N_2}{N'_1}$, as follows:

$$\begin{aligned} \alpha &< 1 + \frac{N_2}{N'_1} - \frac{M_2 (N_1 - M_1)}{(N'_1)^2} \\ \stackrel{(a)}{\Rightarrow} \alpha &< 1 + \frac{N_2}{N'_1} - \frac{N_1 - M_1}{N'_1} \\ \Rightarrow \alpha - \frac{N_2}{N'_1} &< 1 - \frac{N_1 - M_1}{N'_1}, \end{aligned}$$

where (a) holds true because $M_2 \geq N'_1$. Thus, for both (42) and (43) in Lemma 4, only the second sub-case is active. By substituting the value of A_2 from (51) in (42) and (43), it is straightforward to show that P_2 is achievable.

A. Comparison with DoF region

Comparing the delayed CSIT DoF region (35) with the corresponding delayed CSIT GDoF region (48), it is clear that the GDoF region with strong interference ($\alpha > 1$) is always equal to or larger than the corresponding DoF region.

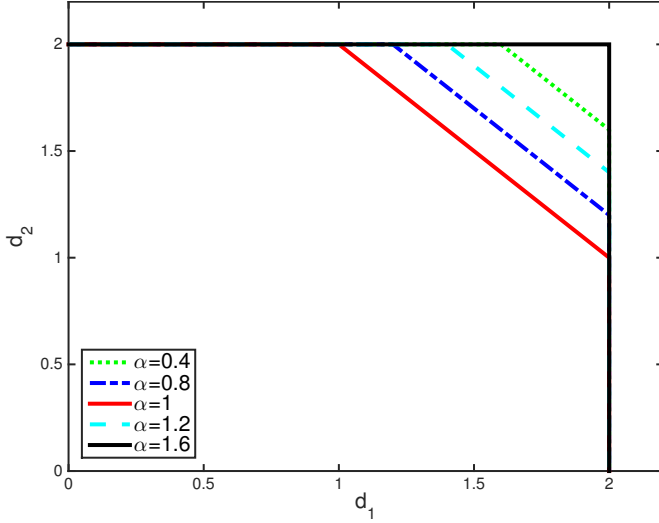


Fig. 4: The GDoF region of the $(2, 2, 3, 2)$ Z-IC at various α .

B. Comparison with perfect CSIT

When $M_2 \leq N_1$ i.e., $N'_1 = M_2$, it is straightforward to show that the delayed CSIT GDoF region shown in (35), coincides with the perfect CSIT GDoF region in (41) when $\alpha > 1$.

When $M_2 > N_1$ and $\alpha > 1$, the non-trivial perfect CSIT bound in (41) is always loose compared to the delayed CSIT bound (48). Thus, when $M_2 > N_1$ and the delayed CSIT bound (35) is active, i.e., when $1 < \alpha < 1 + \frac{N_2}{N_1} - \frac{M_2(N_1 - M_1)}{(N_1)^2}$ (Case II above), the delayed CSIT GDoF region is strictly smaller than the corresponding perfect CSIT GDoF region. Otherwise, for antenna configurations from Case I with $M_2 > N_1$, the delayed CSIT and perfect CSIT GDoF regions coincide for $\alpha > 1$.

C. Sum-GDoF

For Case I above, we see from the shape of the GDoF region in Fig. 2 that the maximum sum-GDoF is given by

$$d_\Sigma = M_1 + N_2.$$

For Case II, the bound (48) is active. The straight line in the (d_1, d_2) plane that defines this bound has slope $-\frac{M_2}{N'_1} \leq -1$, and thus, the maximum sum-GDoF in this case is achieved at the corner point P_1 in Fig. 3, and is equal to

$$d_\Sigma = \left[N_2 + N_1 - \frac{(N_2 + N'_1) N'_1}{M_2} \right] + \frac{(N'_1)^2}{M_2} \alpha$$

which is an increasing linear function of α . This proves the second and remaining part (5) of Corollary 1.

For comparison, we also provide the sum-GDoF with perfect CSIT in the strong interference regime when $M_2 > N_1$ (recall that the GDoF regions for delayed CSIT and perfect CSIT coincide for strong interference when $M_2 \leq N_1$), which, from (41), we find to be as follows:

$$d_\Sigma^p = \begin{cases} M_1 + N_2, & \alpha \geq 1 + \frac{N_2}{N_1} - \frac{M_2 - M_1}{N_1} \\ M_2 + (\alpha - 1) N_1, & \alpha < 1 + \frac{N_2}{N_1} - \frac{M_2 - M_1}{N_1} \end{cases}.$$

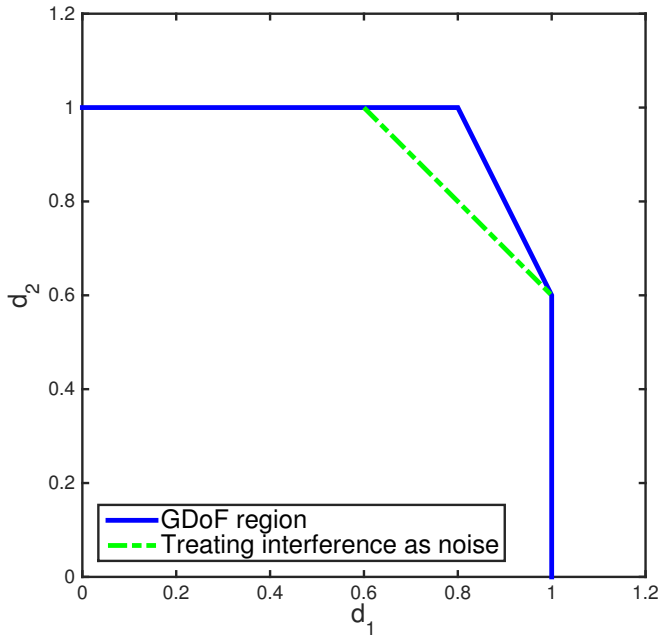


Fig. 5: Sub-optimality of treating interference as noise for the $(1, 2, 1, 1)$ Z-IC when $\alpha = 0.4$.

VIII. DISCUSSION OF RESULTS

A. GDoF vs DoF

When the cross-link of the Z-IC differs in strength from the direct links, this knowledge about the channel statistics is incorporated in the achievability scheme developed in this paper. In such a situation, naively applying the existing DoF-optimal achievability scheme from [3], which incorrectly assumes all three links to be of equal strength, can lead to an achievable region that is strictly sub-optimal. To demonstrate the benefit of incorporating the channel statistics into the achievable scheme through a GDoF analysis, we compared the DoF region ($\alpha = 1$) of the MIMO Z-IC with the corresponding GDoF regions in the weak and strong interference regimes in Sections VI and VII, respectively, and showed that the GDoF region, irrespective of the interference regime, is always larger than or equal to the corresponding DoF region, for all antenna configurations. In general, when the delayed CSIT bound (2) is active, the delayed CSIT GDoF region becomes smaller as α increases from 0 to 1, with the DoF region (at $\alpha = 1$) being the smallest, and then, as α increases for $\alpha > 1$, the GDoF region becomes larger, until the delayed CSIT bound becomes inactive. This is illustrated for the $(2, 2, 3, 2)$ Z-IC in Fig. 4, which also shows that the GDoF regions for $\alpha = 0.4, 0.8$ (weak interference) and $\alpha = 1.2, 1.6$ (strong interference) are strictly larger than the DoF region ($\alpha = 1$).

B. Sub-optimality of treating interference as noise

Our MIMO analysis shows that, unlike the SISO Z-IC with delayed CSIT and weak interference, treating interference as noise (TIN) at R_1 is not in general GDoF-optimal for the MIMO Z-IC with delayed CSIT, even in the weak interference regime. This is illustrated for the $(1, 2, 1, 1)$ MIMO Z-IC with $\alpha = 0.4$ in Fig. 5, where the dotted line shows the achievable GDoF region obtained by treating the interference as noise, which is clearly sub-optimal compared to the actual GDoF region, shown in the figure with solid lines.

C. Delayed CSIT vs perfect CSIT

In Sections VI and VII, we compared the delayed CSIT GDoF region and the corresponding perfect CSIT GDoF region with weak and strong interference, respectively, for various antenna configurations. The insights gained from these comparisons are illustrated below with some representative examples.

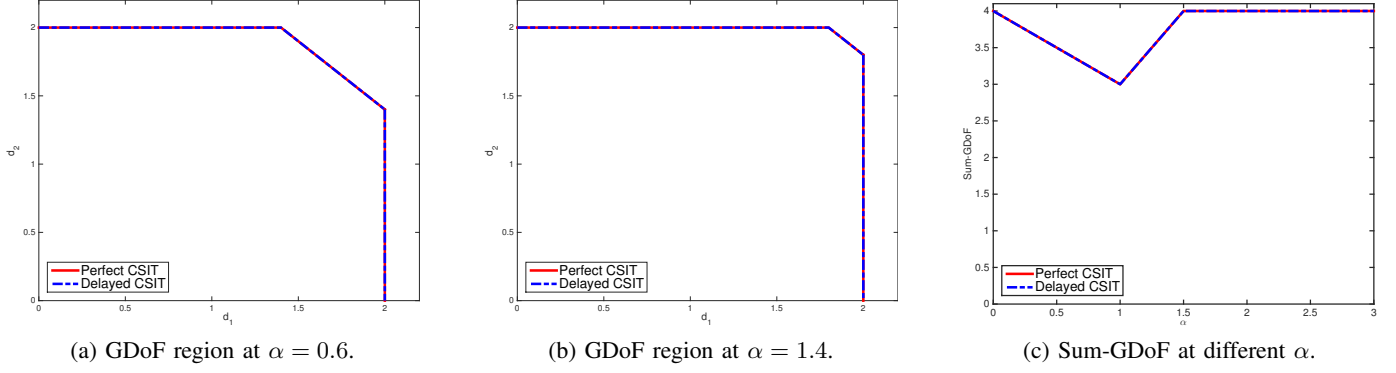


Fig. 6: Comparison of the GDoF region and sum-GDoF of the $(2, 2, 3, 2)$ Z-IC with delayed and perfect CSIT.

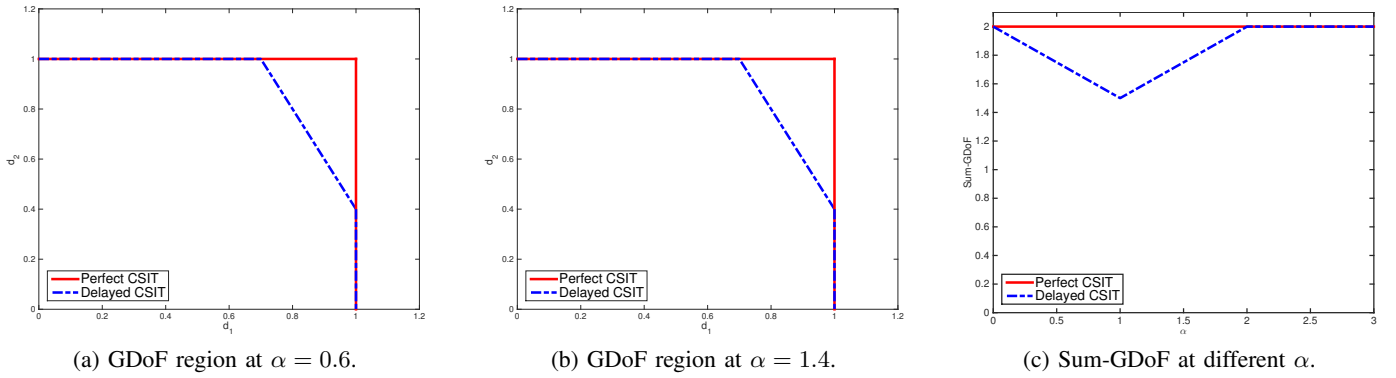


Fig. 7: Comparison of the GDoF region and sum-GDoF of the $(1, 2, 1, 1)$ Z-IC with delayed CSIT and perfect CSIT.

When $M_2 \leq N_1$, delayed CSIT is sufficient to achieve the perfect CSIT GDoF region for all values of α . This is illustrated for the $(2, 2, 3, 2)$ Z-IC in Fig. 6, where the delayed CSIT and perfect CSIT GDoF regions are shown to coincide at $\alpha = 0.6$ (weak interference) and $\alpha = 1.4$ (strong interference). Delayed CSIT is also sufficient to achieve the perfect CSIT GDoF region when $M_2 > N_1$, but only for Case I in Sections VI, i.e., $\alpha \leq 1$ and $\frac{N_1 - M_1}{N_1'} \geq \frac{N_2}{M_2}$, and VII, i.e., $\alpha > 1$ and $\alpha \geq 1 + \frac{N_2}{N_1'} - \frac{M_2(N_1 - M_1)}{(N_1')^2}$.

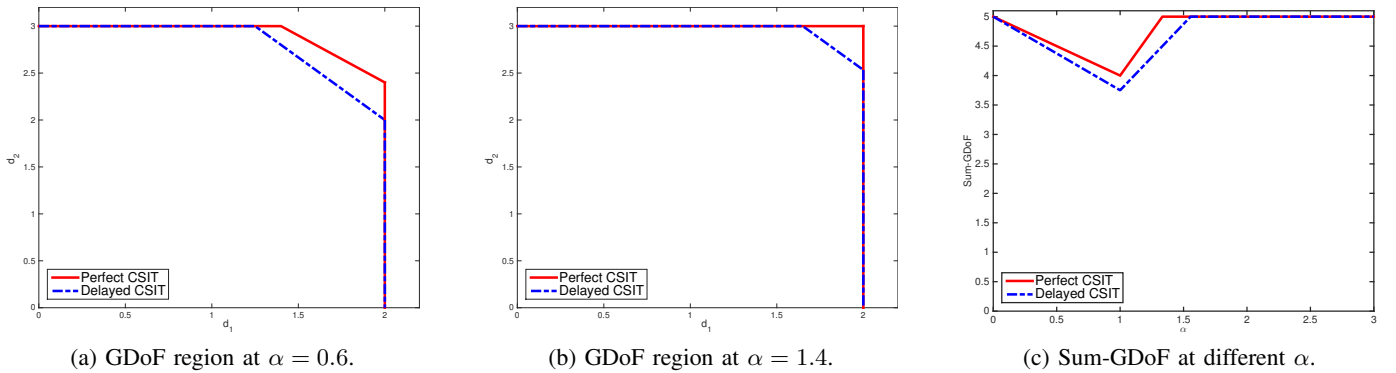


Fig. 8: Comparison of the GDoF region and sum-GDoF of the $(2, 4, 3, 3)$ Z-IC with delayed CSIT and perfect CSIT.

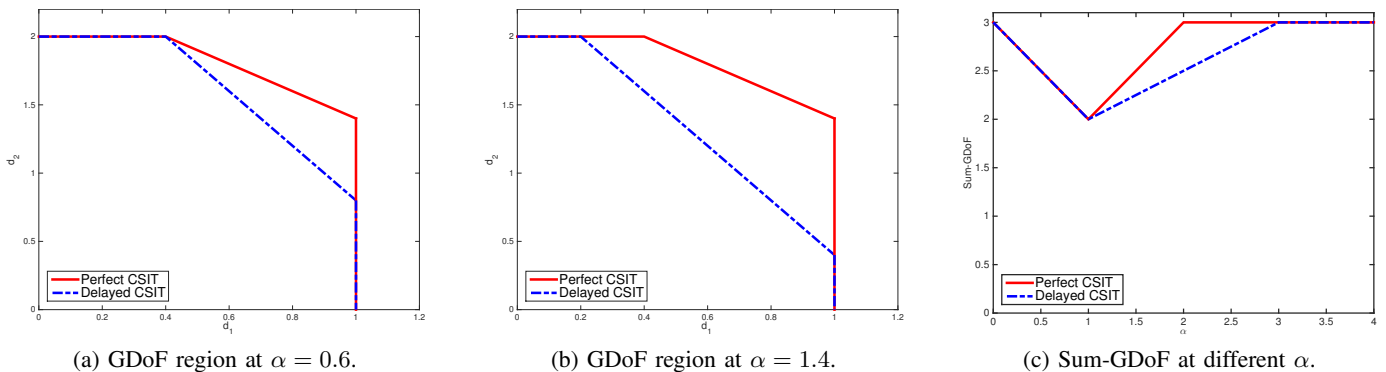


Fig. 9: Comparison of the GDoF region and sum-GDoF of the $(1, 2, 1, 2)$ Z-IC with delayed CSIT and perfect CSIT.

For the remaining choices of antenna tuples and α when $M_2 > N_1$, i.e., Case II of Sections VI and VII, we have already shown that the delayed CSIT GDoF region is strictly smaller than the corresponding perfect CSIT GDoF region (excluding the trivial case of $\alpha = 0$). This is illustrated in Figs. 7 and 8, for the $(1, 2, 1, 1)$ and $(2, 4, 3, 3)$ Z-IC, respectively. For the $(1, 2, 1, 1)$ Z-IC, the perfect CSIT sum-GDoF bound in (41) is never active, and consequently the perfect CSIT GDoF region remains the same at all values of α . For the $(2, 4, 3, 3)$ Z-IC, both the delayed CSIT and perfect CSIT bounds can be active, as seen by the shapes of both the GDoF regions at $\alpha = 0.6$ and $\alpha = 1.4$.

D. Sum-GDoF

In Section VI, we showed that the sum-GDoF is a monotonically decreasing linear function of α in the weak interference regime. For the strong interference regime, it was shown in Section VII that the sum-GDoF is a monotonically increasing linear function of α . In general, this leads to a typical V-shape for the sum-GDoF over the complete range of α . We note that the two segments of this V-shape can have slopes of different magnitudes. In line with our comparison of the DoF vs GDoF regions, the minimum d_Σ is attained at $\alpha = 1$.

From the perfect CSIT sum-GDoF results in Sections VI and VII, it is clear that the perfect CSIT sum-GDoF is also, in general, a V-shaped function of α . We compare the delayed CSIT sum-GDoF with its perfect CSIT counterpart in Figs. 6, 7, 8 and 9, for various antenna configurations. In all of these figures, we see the characteristic V-shape for the delayed CSIT sum-GDoF, with the minimum occurring at $\alpha = 1$. For the $(2, 2, 3, 2)$ Z-IC, the GDoF regions with delayed CSIT and perfect CSIT coincide, and thus, so does the sum-GDoF, as seen in Fig. 6. For the $(1, 2, 1, 1)$ Z-IC, we saw earlier that the perfect CSIT sum-GDoF bound is never active, and this is seen in the constant value of the perfect CSIT sum-GDoF over the range of α in Fig. 7. The sum-GDoF of the $(2, 4, 3, 3)$ Z-IC, in Fig. 8, displays the characteristic V-shape for both the delayed CSIT and perfect CSIT sum-GDoF curves, albeit with different slopes for the two segments of the V-shape in both CSIT regimes. It is clear that for both the $(1, 2, 1, 1)$ and $(2, 4, 3, 3)$ Z-IC, delayed CSIT is not sufficient to attain the perfect CSIT sum-GDoF in the weak interference regime, and for a range of α in the strong interference regime.

An interesting phenomenon is observed for the $(1, 2, 1, 2)$ Z-IC, as seen in Fig. 9. While delayed CSIT is insufficient in this case to achieve the complete perfect CSIT GDoF region in the weak interference regime (except for when $\alpha = 0$), the sum-GDoF plot in Fig. 9c shows that, for weak interference, delayed CSIT is still sufficient to achieve the perfect CSIT sum-GDoF for the $(1, 2, 1, 2)$ Z-IC. The clue to this behavior can be found in Fig. 9a, where we see that, for weak interference ($\alpha = 0.6$ in this case), the perfect CSIT and delayed CSIT GDoF regions with weak interference share the GDoF corner-point (point P_1 in Fig. 3) at which the maximum sum-GDoF is achieved. By analyzing the corner points of the delayed CSIT

and perfect CSIT GDoF regions in the weak interference regime to see where they coincide, it is easy to show that this phenomenon, where delayed CSIT is sufficient to achieve the perfect CSIT sum-GDoF with weak interference in spite of the perfect CSIT GDoF region being strictly larger than its delayed CSIT counterpart, occurs in Case II of Section VI when

$$N_2 = M_2 > N_1.$$

Fig. 9 also shows that, for the $(1, 2, 1, 2)$ Z-IC, delayed CSIT still remains insufficient to achieve the perfect CSIT sum-GDoF in the strong interference regime (when the delayed CSIT bound is active).

IX. CONCLUSION

In this paper, we characterize the GDoF region of the (M_1, M_2, N_1, N_2) MIMO Z-IC under the assumption of delayed CSIT. We obtain a new outer bound by using a combination of a genie, the extremal inequality and maximizing the weighted sum-rate of the two users in the high SNR regime. We next develop a general block-Markov achievability scheme that uses interference quantization to take advantage of the different power levels of the INR and SNR. By specializing this achievability scheme to both weak and strong interference regimes, we show that the outer bound region in each regime is achievable. The GDoF region is found to be equal to or larger than the DoF region over the whole range of α and for all antenna tuples. Moreover, the antenna configurations for which delayed CSIT is sufficient to achieve the perfect CSIT GDoF region or sum-GDoF are characterized. Even in the weak interference regime, we show that treating interference as noise is not a GDoF-optimal strategy in general. To the best of our knowledge, this is the first paper to characterize of the GDoF region of any network with distributed transmitters and delayed CSIT, as well as the first paper to characterize the GDoF region of a MIMO network with delayed CSIT and arbitrary number of antennas at each node. An investigation into the effect of disparate link strengths on the capacity of other MIMO networks, e.g., the 2-user IC, under channel uncertainty, and delayed CSIT in particular, through such GDoF characterization remains an interesting avenue for future research.

APPENDIX A PROOF OF LEMMA 2

To prove Lemma 2, we show that the function $g(r)$ defined in (15) is maximized when $r = 0$. Without any loss of generality, we assume that $M_2 \leq N_1 + N_2$, since it is clear that the function $g(r)$ remains unchanged irrespective of whether $M_2 - r > N_1 + N_2$ or $M_2 - r = N_1 + N_2$. Below, we analyze the strong interference and weak interference cases separately.

Strong interference

We analyze the function $g(r)$ in the strong interference regime ($\alpha > 1$) on a case-by-case basis, as follows:

Case 1) when $M_2 \leq N_1$: $g(r)$ can be written as

$$\frac{f(N_1, (1, M_1), (\alpha, M_2 - r))}{M_2},$$

which is a monotonically decreasing function of r , and is thus maximized at $r = 0$.

Case 2) when $M_2 > N_1$: the first term in $g(r)$ can be written as

$$\begin{aligned} & \frac{f(N_1, (1, M_1), (\alpha, M_2 - r))}{\min(M_2, N_1)} \\ &= \frac{\alpha \min(M_2 - r, N_1) + \min((N_1 - (M_2 - r))^+, M_1)}{N_1}. \end{aligned} \tag{52}$$

Hence, $g(r)$ becomes

$$\frac{\min((N_1 - (M_2 - r))^+, M_1)}{N_1} + \frac{f(M_2 - r, (\alpha, N_1), (1, N_2))}{M_2}. \quad (53)$$

The first term in (53) is monotonically increasing with r , while the second term decreases monotonically with r .

Compared to the full-rank ($r = 0$) case, it can be shown that any increase in the first term, from setting $r > 0$, is always offset by a corresponding decrease in the second term. We illustrate this by analyzing the most optimistic scenario, where $g(r)$ attains its maximum value for $r > 0$, i.e., when $r = M_2$ and $M_1 \geq N_1$. In this scenario, the first term increases by $\frac{N}{N} = 1$ over the full-rank case, while the corresponding loss in the second term is

$$\frac{\alpha N_1 + M_2 - N_1}{M_2} = 1 + (\alpha - 1) \frac{N_1}{M_2}.$$

Thus, the maximum possible increase in $g(r)$, for $r > 0$, compared to the full-rank case is

$$1 - \left[1 + (\alpha - 1) \frac{N_1}{M_2} \right] = -(\alpha - 1) \frac{N_1}{M_2} < 0,$$

proving that $g(r)$ is maximized at $r = 0$.

Weak interference

Since $\alpha \leq 1$, the second term of $g(r)$ can be written as:

$$\frac{f(M_2 - r, (\alpha, N_1), (1, N_2))}{\min(M_2, N_1 + N_2)} = \frac{f(M_2, (\alpha, N_1), (1, N_2))}{\min(M_2, N_1 + N_2)} - l_2(r), \quad (54)$$

where the loss function $l_2(r)$, depends on the antenna tuple as follows (recall that $M_2 \leq N_1 + N_2$):

Case 1) when $M_2 \leq N_2$:

$$l_2(r) = \frac{r}{M_2}.$$

Case 2) when $M_2 > N_2$:

$$l_2(r) = \begin{cases} \frac{r\alpha}{M_2}, & r \leq M_2 - N_2 \\ \frac{(M_2 - N_2)\alpha + (N_2 - (M_2 - r))}{M_2}, & r > M_2 - N_2. \end{cases}$$

The third term of $g(r)$ can also be written as

$$\frac{\alpha \min(M_2 - r, N_1)}{\min(M_2, N_1)} = \alpha - l_3(r), \quad (55)$$

where the loss function $l_3(r)$ depends on the antenna tuple as follows:

Case 1) when $M_2 \leq N_1$:

$$l_3(r) = \frac{\alpha r}{M_2}.$$

Case 2) when $M_2 > N_1$:

$$l_3(r) = \begin{cases} 0, & r \leq M_2 - N_1 \\ \frac{N_1 - (M_2 - r)}{N_1} \alpha, & r > M_2 - N_1. \end{cases}$$

Now, by substituting (54) and (55), we can write $g(r)$ as

$$x(r) + [l_3(r) - l_2(r)],$$

where

$$x(r) = \frac{f(N_1, (1, M_1), (\alpha, M_2 - r))}{N_1} + \frac{f(M_2, (\alpha, N_1), (1, N_2))}{M_2} - \alpha$$

is a monotonically decreasing function of r . Thus, to prove that the function $g(r)$ is a monotonically decreasing function of r , which consequently attains its maximum at $r = 0$, it suffices to prove that

$$l_3(r) - l_2(r) \leq 0, \quad (56)$$

for $0 \leq r \leq M_2$. We prove (56) below on a case-by-case basis, for all antenna tuples.

Case 1) when $M_2 \leq N_1$, $M_2 \leq N_2$:

$$l_3(r) - l_2(r) = (\alpha - 1) \frac{r}{M_2} \leq 0.$$

Case 2) when $N_2 < M_2 \leq N_1$:

$$\begin{aligned} l_3(r) - l_2(r) &= \begin{cases} (\alpha - 1) \frac{r}{M_2}, & r \leq M_2 - N_2 \\ (\alpha - 1) \frac{N_2 - (M_2 - r)}{M_2}, & r > M_2 - N_2 \end{cases} \\ &\leq 0. \end{aligned}$$

Case 3) when $N_1 < M_2 \leq N_2$: In this case, there are two sub-cases, which we analyze separately below.

When $r \leq M_2 - N_1$, we have

$$l_3(r) - l_2(r) = 0 - \frac{r}{M_2} \leq 0.$$

When $r > M_2 - N_1$, we have

$$\begin{aligned} l_3(r) - l_2(r) &= \frac{(N_1 - M_2 + r)\alpha}{N_1} - \frac{r}{M_2} \\ &= \left(1 - \frac{M_2}{N_1}\right)\alpha + r\left(\frac{\alpha}{N_1} - \frac{1}{M_2}\right) \\ &\stackrel{(a)}{\leq} \left(1 - \frac{M_2}{N_1}\right)\alpha + M_2\left(\frac{\alpha}{N_1} - \frac{1}{M_2}\right)^+ \\ &\stackrel{(b)}{\leq} 0, \end{aligned}$$

where we have used $r \leq M_2$ in (a) and $N_1 < M_2$ in (b).

Case 4) when $N_1 < M_2$, $N_2 < M_2$: We have three sub-cases, which are analyzed separately below.

When $r \leq M_2 - N_1$, we have

$$l_3(r) - l_2(r) = 0 - l_2(r) \leq 0.$$

The second sub-case is defined as $M_2 - N_1 < r \leq M_2 - N_2$, and we obtain

$$\begin{aligned} l_3(r) - l_2(r) &= \frac{(N_1 - M_2 + r)\alpha}{N_1} - \frac{r\alpha}{M_2} \\ &= \left(1 - \frac{M_2}{N_1}\right)\left(1 - \frac{r}{M_2}\right)\alpha \\ &\leq 0. \end{aligned}$$

Thirdly, when $r > M_2 - N_1$ and $r > M_2 - N_2$, we have

$$\begin{aligned}
& l_3(r) - l_2(r) \\
&= \frac{(N_1 - M_2 + r)\alpha}{N_1} - \frac{(M_2 - N_2)\alpha + (N_2 - M_2 + r)}{M_2} \\
&= \left(\frac{M_2 - N_2 - r}{M_2} \right) - \alpha \left(\frac{M_2 - r}{N_1} - \frac{N_2}{M_2} \right) \\
&\stackrel{(a)}{\leq} \left(\frac{M_2 - N_2 - r}{M_2} \right) - \alpha \left(\frac{M_2 - r}{M_2} - \frac{N_2}{M_2} \right) \\
&= \left(\frac{M_2 - N_2 - r}{M_2} \right) (1 - \alpha) \\
&\stackrel{(b)}{\leq} 0,
\end{aligned}$$

where we have used $N_1 < M_2$ in (a) and $M_2 - N_2 - r < 0$ in (b).

Thus, for all antenna tuples, we have shown that the function $g(r)$ is a monotonically decreasing function of r in the weak interference regime, and is hence maximized at $r = 0$.

APPENDIX B ACHIEVABLE GDOF REGION

In this appendix, we show that the achievable GDoF region for the MIMO Z-IC using the general achievability scheme from Section V is given by (20)-(24). The 2-user MACs obtained in (18) and (19) are shown below (we omit the block index b for the channel matrices):

$$\begin{aligned}
\underbrace{y_1[b] - \eta_b}_{Y'_1} &= \underbrace{\sqrt{\rho^\alpha} H_{12} x_{2c}(l_{b-1})}_{X'_c} + \underbrace{\sqrt{\rho} H_{11} u_1(w_{1,b})}_{X'_1} + Z_1, \\
\underbrace{\begin{bmatrix} y_2[b] \\ \eta_b \end{bmatrix}}_{Y'_2} &= \begin{bmatrix} \sqrt{\rho} H_{22} \\ 0 \end{bmatrix} \underbrace{x_{2c}(l_{b-1})}_{X'_c} + \begin{bmatrix} \sqrt{\rho} H_{22} \\ \sqrt{\rho^\alpha} H_{12} \end{bmatrix} \underbrace{u_2(w_{2,b})}_{X'_2} + Z_2,
\end{aligned}$$

where, Z_i is the AWGN at receiver R_i , and, for simplicity, we have also defined Y'_1, Y'_2, X'_c, X'_1 and X'_2 , such that X'_c, X'_1 and X'_2 have achievable rates \bar{R}_c, \bar{R}_1 and \bar{R}_2 , respectively, and the associated GDoF are respectively d_η, d_{1b} and d_{2b} (from (20)-(24)). We note that $X'_c \sim \mathcal{CN}(0, Q_c)$, $X'_1 \sim \mathcal{CN}(0, Q_1)$ and $X'_2 \sim \mathcal{CN}(0, Q_2)$ are independent, with covariance matrices $Q_c \triangleq \mathbf{I}_{M_2}$, $Q_1 \triangleq \mathbf{I}_{M_1}$ and $Q_2 \triangleq \rho^{-A_2} \mathbf{I}_{M_2}$. From [25], the achievable rate region for the 2-user MAC at receiver $R_i, i \in \{1, 2\}$ is given by the following information-theoretic inequalities :

$$\bar{R}_c \leq I(X'_c; Y'_i | X'_i, \mathcal{H}^n), \quad (57)$$

$$\bar{R}_i \leq I(X'_i; Y'_i | X'_c, \mathcal{H}^n), \quad (58)$$

$$\bar{R}_c + \bar{R}_i \leq I(X'_c, X'_i; Y'_i | \mathcal{H}^n). \quad (59)$$

From (57), we obtain the following GDoF bounds:

$$\begin{aligned}
\bar{R}_c &\leq I(X'_c; Y'_1 | X'_1, \mathcal{H}^n) \\
&= h(Y'_1 | X'_1, \mathcal{H}^n) - h(Y'_1 | X'_1, X'_c, \mathcal{H}^n) \\
&= \log \left| \mathbf{I}_{N_1} + \rho^\alpha H_{12} Q_c H_{12}^\dagger \right| + \mathcal{O}(1) \\
&\stackrel{(a)}{=} \alpha \min(M_2, N_1) \log \rho + \mathcal{O}(1) \\
\Rightarrow d_\eta &\leq \alpha N'_1,
\end{aligned} \quad (60)$$

where we have substituted $Q_c = \mathbf{I}_{M_2}$, (a) follows from Lemma 1, and finally dividing both sides by $\log \rho$ as $\rho \rightarrow \infty$ gives the GDoF bound (60). Similarly, we obtain

$$\begin{aligned}
\bar{R}_c &\leq I(X'_c; Y'_2 | X'_2, \mathcal{H}^n) \\
&= \log \left| \mathbf{I}_{N_2} + H_{22} Q_c H_{22}^\dagger \right| + \mathcal{O}(1) \\
&\stackrel{(a)}{=} \min(M_2, N_2) \log \rho + \mathcal{O}(1) \\
&\stackrel{(b)}{\Rightarrow} d_\eta \leq N_2,
\end{aligned} \tag{61}$$

where in (a), we have substituted $Q_c = \mathbf{I}_{M_2}$ and used Lemma 1, and (b) uses the antenna assumption $N_2 \leq M_2$ from (3). Combining (60) and (61), we obtain (20).

The bound (21) for d_{1b} follows from (58), with $i = 1$, in a similarly straightforward manner. Bound (22) is obtained from (59), with $i = 1$, by a direct application of Lemma 1, as shown below:

$$\begin{aligned}
&\bar{R}_c + \bar{R}_1 \\
&\leq I(X'_c, X'_1; Y'_1) \\
&= \log \left| \mathbf{I}_{N_1} + \rho^\alpha H_{12} Q_c H_{12}^\dagger + \rho H_{11} Q_1 H_{11}^\dagger \right| + \mathcal{O}(1) \\
&= f(N_1, (\alpha, M_2), (1, M_1)) \log \rho + \mathcal{O}(1),
\end{aligned}$$

where we have substituted $Q_c = \mathbf{I}_{M_2}$ and $Q_1 = \mathbf{I}_{M_1}$, and dividing both sides by $\log \rho$ gives the required bound (22).

To prove (23), we use the singular value decomposition (SVD) of $\begin{bmatrix} H_{22} \\ H_{12} \end{bmatrix} = U \Lambda V^\dagger$, where $U \in \mathbb{C}^{(N_1+N_2) \times M_2}$ is a matrix such that $U^\dagger U = \mathbf{I}_{M_2}$, Λ is a $M_2 \times M_2$ diagonal matrix such that the diagonal elements consist of the singular values of $\begin{bmatrix} H_{22} \\ H_{12} \end{bmatrix}$ and $V \in \mathbb{C}^{M_2 \times M_2}$ is matrix such that $V^\dagger V = \mathbf{I}_{M_2}$. Using (58), we obtain (23) as follows (we have omitted the $\mathcal{O}(1)$ terms):

$$\begin{aligned}
\bar{R}_2 &\leq I(X'_2; Y'_2 | X'_c, \mathcal{H}^n) \\
&= \log \left| \mathbf{I}_{N_2+N_1} + \begin{bmatrix} \sqrt{\rho} H_{22} \\ \sqrt{\rho^\alpha} H_{12} \end{bmatrix} Q_2 \begin{bmatrix} \sqrt{\rho} H_{22}^\dagger & \sqrt{\rho^\alpha} H_{12}^\dagger \end{bmatrix} \right| \\
&\stackrel{(a)}{=} \log \left| \mathbf{I} + \begin{bmatrix} \rho^{(1-A_2)} I_{N_2} & 0 \\ 0 & \rho^{(\alpha-A_2)} I_{N_1} \end{bmatrix} \begin{bmatrix} H_{22} \\ H_{12} \end{bmatrix} \begin{bmatrix} H_{22}^\dagger & H_{12}^\dagger \end{bmatrix} \right| \\
&\stackrel{(b)}{=} \log \left| \mathbf{I} + \begin{bmatrix} \rho^{(1-A_2)} I_{N_2} & 0 \\ 0 & \rho^{(\alpha-A_2)} I_{N_1} \end{bmatrix} U \Lambda \underbrace{V^\dagger V}_{\mathbf{I}_{M_2}} \Lambda U^\dagger \right| \\
&\stackrel{(c)}{=} \log \left| \mathbf{I} + \begin{bmatrix} \rho^{(1-A_2)} U_2 \Lambda \\ \rho^{(\alpha-A_2)} U_1 \Lambda \end{bmatrix} \begin{bmatrix} \Lambda U_2^\dagger & \Lambda U_1^\dagger \end{bmatrix} \right| \\
&\stackrel{(d)}{=} \log \left| \mathbf{I}_{M_2} + \begin{bmatrix} \tilde{U}_2^\dagger & \tilde{U}_1^\dagger \end{bmatrix} \begin{bmatrix} \rho^{(1-A_2)} \tilde{U}_2 \\ \rho^{(\alpha-A_2)} \tilde{U}_1 \end{bmatrix} \right| \\
&= \log \left| \mathbf{I}_{M_2} + \rho^{(1-A_2)} \tilde{U}_2^\dagger \tilde{U}_2 + \rho^{(\alpha-A_2)} \tilde{U}_1^\dagger \tilde{U}_1 \right| \\
&\stackrel{(e)}{=} f(M_2, (1-A_2, N_2), (\alpha-A_2, N_1)) \log \rho
\end{aligned}$$

and dividing both sides by $\log \rho$ as $\rho \rightarrow \infty$, we obtain (23). In the above derivation, (a) is obtained by substituting $Q_2 = \rho^{-A_2} \mathbf{I}_{M_2}$, (b) uses the SVD of $\begin{bmatrix} H_{22} \\ H_{12} \end{bmatrix}$ defined earlier, (c) uses the following partitioning, $U = \begin{bmatrix} U_2 \\ U_1 \end{bmatrix}$, where $U_i \in \mathbb{C}^{N_i \times M_2}$, for $i \in \{1, 2\}$, and finally (d) follows, after defining

$\tilde{U}_i \triangleq U_i \Lambda$, from the identity $|\mathbf{I} + AB| = |\mathbf{I} + BA|$. The asymptotic approximation in (e) is a direct consequence of Lemma 1.

To obtain (24), we proceed from (59) as follows (the conditioning on \mathcal{H}^n and the $\mathcal{O}(1)$ terms are omitted):

$$\begin{aligned}
& \bar{R}_c + \bar{R}_2 \\
& \leq I(X'_c, X'_2; Y'_2) \\
& = h(Y'_2) + \mathcal{O}(1) \\
& = h(\sqrt{\rho^\alpha} H_{12} X'_2) + h(\sqrt{\rho} H_{22} (X'_c + X'_2) | \sqrt{\rho^\alpha} H_{12} X'_2) \\
& \stackrel{(a)}{\geq} (\alpha - A_2)^+ N'_1 \log \rho \\
& \quad + h(\sqrt{\rho} H_{22} (X'_c + X'_2) | \sqrt{\rho^\alpha} H_{12} X'_2, X'_2) \\
& = (\alpha - A_2)^+ N'_1 \log \rho + h(\sqrt{\rho} H_{22} X'_c) \\
& = (\alpha - A_2)^+ N'_1 \log \rho + N_2 \log \rho,
\end{aligned}$$

where (a) is true because conditioning reduces entropy, and any region contained in the achievable region is also achievable. By dividing both sides above by $\log \rho$ as $\rho \rightarrow \infty$, we obtain (24).

REFERENCES

- [1] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inform. Th.*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.
- [2] C. Vaze and M. Varanasi, "The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSIT," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, Aug. 2011, pp. 199–203.
- [3] K. Mohanty and M. Varanasi, "Degrees of freedom of the MIMO Z-interference channel with delayed CSIT," *Communications Letters, IEEE*, vol. 19, no. 12, pp. 2282–2285, Dec 2015.
- [4] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Trans. Inform. Th.*, vol. 58, no. 7, pp. 4396–4417, Jul. 2012.
- [5] M. Abdoli, A. Ghasemi, and A. Khandani, "On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT," *Information Theory, IEEE Transactions on*, vol. 59, no. 10, pp. 6542–6561, 2013.
- [6] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," in *ISIT*, Aug. 2011, pp. 209–213.
- [7] D. T. H. Kao and A. S. Avestimehr, "Linear degrees of freedom of the MIMO X-channel with delayed CSIT," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 366–370.
- [8] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *Information Theory, IEEE Transactions on*, vol. 54, no. 12, pp. 5534–5562, Dec 2008.
- [9] S. Karmakar and M. Varanasi, "The generalized degrees of freedom region of the MIMO interference channel and its achievability," *Information Theory, IEEE Transactions on*, vol. 58, no. 12, pp. 7188–7203, Dec 2012.
- [10] C. Vaze, S. Karmakar, and M. Varanasi, "On the generalized degrees of freedom region of the MIMO interference channel with no CSIT," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, July 2011, pp. 757–761.
- [11] S. Karmakar and M. Varanasi, "The generalized multiplexing gain region of the slow fading MIMO interference channel and its achievability with limited feedback," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, July 2012, pp. 3135–3139.
- [12] J. Chen, P. Elia, and S. Jafar, "On the two-user MISO broadcast channel with alternating CSIT: A topological perspective," *Information Theory, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
- [13] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *Information Theory, IEEE Transactions on*, vol. 53, no. 5, pp. 1839–1851, May 2007.
- [14] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally correlated MIMO networks with delayed CSIT," *Information Theory, IEEE Transactions on*, vol. 60, no. 1, pp. 494–514, Jan 2014.
- [15] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep 1979.
- [16] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, Jul 1984.
- [17] C. Suh and D. N. C. Tse, "Feedback capacity of the gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.
- [18] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *Information Theory, IEEE Transactions on*, vol. 59, no. 1, pp. 315–328, Jan. 2013.
- [19] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with Mixed CSIT," *Communications Letters, IEEE*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.

- [20] J. Chen and P. Elia, “Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT,” in *Proc. Inf. Theory and App. Workshop (ITA)*, Feb. 2013.
- [21] K. Mohanty and M. Varanasi, “Degrees of freedom of the MIMO Z-interference channel with mixed CSIT,” *Communications Letters, IEEE*, accepted for publication.
- [22] C. S. Vaze and M. K. Varanasi, “The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT,” *IEEE Trans. Inform. Th.*, vol. 58, no. 8, pp. 5354–5374, Aug. 2012.
- [23] H. Weingarten, Y. Steinberg, and S. Shamai, “The capacity region of the gaussian multiple-input multiple-output broadcast channel,” *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [24] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.
- [25] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2012.